DOCUMENT RESUME

ED 033 484 EC 004 358

AUTHOR	Budoff, M.
TITLE	An Educational Test of the Learning Potential Hypothesis with Adolescent Mentally Retarded Special Class Children, Volume I. Final Report.
INSTITUTION	Cambridge Mental Health Center, Mass.
Spons Agency	Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No	BR-6-1184
Pub Date	[69]
Grant	OEG-32-31-0000-6019
Note	242p.
Available from	Cambridge Mental Health Center, 20 Sacramento Street, Cambridge, Massachusetts 02138.
EDRS Price	EDRS Price MF-$1.00 HC Not Available from EDRS.
Descriptors	Ability, *Academic Achievement, Average Students, Educable Mentally Handicapped, Electricity, Evaluation Methods, *Exceptional Child Research, *Instructional Materials, Junior High School Students, Manipulative Materials, *Mentally Handicapped, Nonverbal Ability, Nonverbal Learning, Sciences, Student Evaluation, *Teaching Methods

Abstract

To test the hypothesis that the amount learned from a manipulative, nonverbal oriented unit on electricity would be better predicted by a learning potential assessment procedure than from an IQ based classification (educable mentally retarded status), an EMR group and a regular class non EMR group were both exposed to the unit; untaught EMRs served as controls. Students were assessed on pretests; students who were high scorers or gainers were considered to have strong learning potential. Results showed that both taught groups knew more about electricity after the unit than before, and that EMR gainers and highscorers learned more than nongainers. A second study showed no difference in amount learned between the unit and a lecture-demonstration unit. Conclusions were as follow: certain EMR's are very capable when taught by nonverbal material; high scorers in the EMR range and low achieving regular class children may require novel presentations that will minimize effects of poor literacy skills; and EMR high scorers might do better in a regular class if curriculum changes were made to compensate for their reading deficits or if they were given extensive work

BR-61184-VOL I
?A-40
OE-BR

FINAL REPORT
Project No. 6-1184
Grant No. 32-31-0000-6019

Volume I

AN EDUCATIONAL TEST OF THE LEARNING POTENTIAL HYPOTHESIS WITH ADOLESCENT MENTALLY RETARDED SPECIAL CLASS CHILDREN

M. Budoff
Principal Investigator

J. Meskin
Educational Coordinator

Cambridge Mental Health Center
20 Sacramento Street
Cambridge, Massachusetts, 02138

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research
Division of Handicapped Children and Youth

Summary of Final Report

of

AN EDUCATIONAL TEST OF THE LEARNING POTENTIAL HYPOTHESIS WITH ADOLESCENT MENTALLY RETARDED SPECIAL CLASS CHILDREN

M. Budoff and J. Meskin

Research Institute for Educational Problems
Harvard Medical School

Project No. 6-1184
Bureau of Research
Division of Handicapped
Children and Youth
U.S. Office of Education

20 Sacramento Street
Cambridge, Massachusetts 02138

CONTENTS

APPENDICES

VOLUME II

LIST OF FIGURES

LIST OF TABLES

Page

Page

This report is dedicated to the memory of Miss Eleanor Maloy, Head of Special Classes in the Waltham, Massachusetts, School System. Her inspiration and encouragement did much to further this line of research.

M. Budoff

J. Meskin

Acknowledgements

Many staff members helped conceptualize this project and facilitated its progress, but five colleagues were especially instrumental to its successful conclusion. We would like to thank especially: Robert H. Harrison, Ph.D., Rosalind Folman, M.A., Deborah Glaser Kemler, B.A., and Linda Ronn, B.A., who made significant contributions in formulating various sections of the project and were separately responsible for various portions of the data analysis; and Mrs. Karen Corbett Howell, M.A.T., who acted as teacher-project assistant during the major study and wrote the major portion of the lecture-demonstration curriculum.

Without the invaluable aid and assistance of the Waltham and Boston School Departments, we would not have been able to conduct this project at all. The late Miss Eleanor Maloy, Head of the Special Classes Department in Waltham, encouraged us to conduct our pilot classes in the Royal E. Robbins School. Due to the excellent cooperation and continuing interest of Dr. Vincent P. Connors, Head of the Special Class Department in the Boston Schools, and the late Mr. Charles O. Ruddy, Associate Superintendent for Pupil Services, we were able to steer around the inevitable difficulties involved in doing research of this kind in the crowded and complex administrative problems of an on-going school. Our especial thanks must go to the school personnel who both educated us about these problems, and tolerated our innocence. Miss Frances Mullen, teacher, Mr. George Sawyer, Principal, Patrick F. Gavin Jr. High School, and Mr. John Whelan, Principal, Donald McKay School did much to facilitate the studies within their respective schools.

Of course, this work would have been impossible without the concurrence of Mr. John Fitzgerald, Superintendent of the Waltham Public Schools, and the Board of Superintendents and the School Committee of Boston.

Finally, special thanks is due to our project teachers, Mrs. Jean H. Rosner and Mr. Robert King for both their enthusiasm and effort in the classroom and their cooperation with the often exacting demands of a research study.

Milton Budoff, Ph.D.

Joan Meskin, M.A.T.

SUMMARY OF FINDINGS FROM A STUDY OF "THE EDUCATIONAL IMPLICATIONS OF THE LEARNING POTENTIAL HYPOTHESIS WITH MENTALLY RETARDED ADOLESCENTS"

This research program has been investigating whether the ability displayed by children in special classes or institutions for the retarded on a novel assessment strategy would also predict their learning competence on specially designed curricula. The novel assessment strategy employed presents a non-verbal reasoning task, the Kohs Block Design problems, in a format that permits the student to learn how to solve the problem and then to demonstrate how well he has learned by requiring him to transfer this understanding to non-coached problems. More specifically, the procedure involves a pretest administration of the block designs, followed by a coaching session in which the child is given a directed learning experience oriented toward teaching him how one solves simple and complex block design problems. Within a day or two and again one month following this coaching session the original series of block designs is readministered.

Three patterns of response have been observed. One group of children tend to show no need for the coaching since they solved quite complex problems on the initial administration. We have called these children <u>high scorers</u>. They seem able to learn to do the problems successfully as they progress through the test, in much the same manner as children of higher IQ probably learn from their experiences to solve these problems successfully. The second group of children shows a marked increase in competence on the post coaching administrations and we have called this group <u>gainers</u>. We have assumed that this group has learned from the tutorial experience and thus improved their performance following the coaching session. A third group of persons are essentially nonlearners in that they do not markedly improve their scores following coaching and we have called this group <u>nongainers</u>.

The present summary reports a study which tested the educational implications of this difference in ability displayed within a homogeneous IQ group of educable mental retardates (IQ range was generally between 60 and 80, the generally accepted IQ range for this group). A variety of psychological tests and learning tasks have demonstrated that high scorers and gainers tend to learn more rapidly, tend to show more ability to reason on non-verbal kinds of problems than the nongainers. The hypothesis has been advanced in other reports from the learning potential project that gainers and high scorers may more truly be considered educationally retarded. Nongainers may be considered more intrinsically retarded, functioning in a manner similar to that described for retardates. Some supporting evidence for this conclusion is evident in the fact that in at least one study in which a comparison was made (Budoff and Pagell, 1968) using mental age and chronological age controls, the nongainers' performance tended to parallel and be somewhat inferior to their mental age peers.

A test of the educational implications of these findings was made through presentation of a specially designed curriculum unit which permitted learning to take place in a non-verbal, manipulative mode of instruction. The curriculum was so designed that student success depended minimally on the child's ability to read or to verbalize satisfactorily. While some recording of data that the children observed was required, at no point did success in the course depend on the quality of his written work. The hypothesis of the study, in brief, was that high scorers and gainers would perform similarly to chronological age controls, and learn markedly more than their nongainer special classmates.

A unit which teaches some simple concepts of electricity and was developed by the Elementary Science Study of the Educational Development Center, Newton, Massachusetts (Batteries and Bulbs) was adapted for this study. It was modified so that it was somewhat more structured, and presentation of materials was controlled so that evaluation of the learning could take place. A pilot attempt to present "discovery" type of science materials in the recommended unstructured way indicated that the special class children were overwhelmed by the lack of structure. They tended not to be able to focus satisfactorily on the rich possibilities of exploration and discovery presented to them. It seemed to us that they needed aid, structure and guidance at least in beginning this type of learning process. During the course of the presentation of this unit the teacher naturally sought to have them explore as much as possible within the given topic units. The topic units were prescribed in advance in an effort to make the learning experience as equivalent as possible across all classes.

An evaluation instrument was designed by the project staff which sought a means of testing the children's comprehension without at the same time having the measure contaminated by their inability to read or to express themselves satisfactorily. The solution arrived at was to present life sized models of actual circuits on masonite pegboards and use a multiple choice, minimal response strategy to test their understanding of what was going on in a given circuit. A second section of the evaluation instrument required the children to make similar judgements in a diagram format in which two thirds of the set-ups were drawn as they would appear in reality and in life size. The last third presented diagrams in a schematic textbook form which they were required to master in the course, but which was also taught to them at the time of each of the testing sessions so that "secret language" would not serve as an obstacle to comprehension. The test sought to determine whether the understanding that might be most easily evident on the boards format when the materials are presented in life size concrete object form would also be manifested in diagrams presented in life size form, and in the more abstract schematic, "secret language" format.

Thirty-three children each from special and regular classes were enrolled in six classes of 12 children each (one child in each of the six classes had to be eliminated eventually because of excessive absence) were taught by one teacher especially trained in presenting the electricity unit. All the taught children were students at one junior high school. A non-taught control group of special class students from the same junior high school and another junior high school was matched to the children in the classes by IQ, chronological age and learning potential status. The course was given twice a week for a one hour session for a period of 13 weeks. All groups missed two classes because of "snow" days making a total of 24 classes. All the classes had similar lessons and progressed at a similar rate through the same materials by the end of the unit. The evaluation instrument was administered prior to the start of classes and in the week following the end of classes.

The results agreed with the hypotheses of the study. Teaching led to greater understanding of electricity. Gainer and high scorer EMRs learned more in the electricity course than nongainer EMRs. When the performance of the taught special class group was compared with the scores obtained by the regular class children who had received the course, the regular class children had higher scores, and somewhat fewer differences on the posttest. When the criterion used was improvement in comprehension (i.e., posttest score minus pretest score) there was no difference between the scores obtained by special or regular class children. Both groups tended to gain equally from the classes, but because of the higher starting level of the regular class children, their posttest scores tended to be higher. Thus, the major discriminating effect was learning potential status. That is, regardless of special or regular class placement, high scorers tended to have learned most from the course, gainers next and nongainers learned least. These results are consistent with the findings of our other studies.

Of the 33 regular class children, 22 fell into a low achieving status group, that is, they had obtained an average grade point average of less than 2.0 (C) during the first four of five marking periods of the school year in which the classes were taught. It was hypothesized that gainers and high scorers might be educationally rather than mentally retarded children since their educational difficulties revolve in large part around their failure to acquire the skills in the language arts and reading. Thus, while their poor scholastic aptitude scores (their low IQ) is a correct prediction of their school status, it may be an incorrect statement of their prognosis, if a different educational treatment were accorded them. To test this hypothesis in the context of this study, the scores on the electricity evaluation test were reanalyzed comparing the taught special class and low achieving regular class children. The results from these analyses more

than support this hypothesis. That is, on the pretest scores, the regular class children score somewhat higher as might be expected. However, with the administration of the unit, this difference disappeared, as evidenced by the lack of difference on scores on the posttest and improvement scores (posttest minus pretest scores).

In both these latter analyses, the major differentiating effect was learning potential status. On the posttest scores this was true for every single subsection of the evaluation instrument. That is the children who demonstrated their ability to profit from the learning potential procedure and were categorized either as high scorers or gainers tended to learn markedly more in the unit than those who were nongainers regardless of whether they were in the special or regular class.

The differences in the ability to profit from the electricity course displayed by the special and regular class students differing in learning potential status was also tested by examining their ability to generalize a correct answer from one mode of presentation to another. If a subject learned the underlying concept involved in any set of matched questions, he should be able to respond to it correctly regardless of the mode of presentation. The subject who responded correctly to an item in only one mode, and not in the others, was presumed to be responding on the basis of the specific item, and not on the basis of the more general concept, demonstrating that he had either not learned the concept, or was unable to apply the concept in other instances of the same problem. Thus another test of the differences in learning displayed by the students in the electricity curriculum classes was their consistency of response across items in a set. Consistency of response was scored when the subject's correct responses were viewed as a percentage of the total possible opportunities to be consistent (Score #1), or as a percentage of those items in the sets in which the student correctly responded to one item (Score #2).

The first score, which was based on the total number of times it was theoretically possible to be consistently correct, can be considered a measure of both S's knowledge of the course work and of his ability to generalize the same concept across different presentations. The greater his knowledge of the material, the greater his chances to be consistently correct. The second score excluded those sets on which S was completely wrong, i.e., where he had no knowledge of the material in question, and included only those sets on which Ss demonstrated knowledge. This score more clearly presents the generalizing inclination of the student, since it is a measure of the consistency of the student on the sets on which he displayed some understanding.

The scoring for consistent application of a concept or principle yields findings similar to the comparisons of the absolute level of evaluation test scores. Teaching the unit improved consistency of response, when knowledge of electricity (Score #1) was considered, and when the consistency score was based only on the sets in which the student demonstrated some understanding (Score #2). While regular class students tended to respond more consistently on the pretest, consistent responses to the posttest were more critically a function of learning potential status. Learning potential status was the variable, singly or in combination with the pre- versus post scores, that accounted for most remaining differences in performance among special and regular class students. High scorers demonstrated the best ability to generalize taught concepts; gainers demonstrated the next best ability, nongainers the least. But while differences in learning potential status did not result in large differences in pre- or posttest scores among regular class students, the differences were very large among the special class students, especially after teaching. Typically, even the regular class nongainer performed relatively like his gainer and high scorer peers, i.e., they tended to show appreciable gains in consistency following teaching. By contrast, the special class nongainers showed little change in consistency following teaching even when they had correctly answered one item in a set and their broader lack of knowledge was not considered (Score #2), while the special class gainers and high scorers demonstrated marked improvement, especially following teaching. Even without being taught, gainers and high scorers improved their consistency scores somewhat. In conclusion, then, when posttest or gain scores were considered, differences between special and regular class students disappeared, and differences by learning potential status became the main differentiating variable.

The results seemed to suggest that exploratory-manipulative materials were most appropriate for tapping the latent ability of the high scorer and gainer EMRs, because of their handicaps in the language arts areas. To test this conclusion, a parallel lecture-demonstration version of the electricity curriculum was written; two classes from a junior high school were administered the two units. Care was taken to exclude all student manipulation experience from the lecture-demonstration presentation, but not penalize their performance on the evaluation instrument, to which a verbal reasons section was added. No differences between performances on the two units was evident. Only learning potential status differentiated among the students in both special class presentations. When the electricity evaluation test scores were adjusted for the differences in IQs, this effect became markedly attenuated. The discussion of this pilot study concludes that a good unit, well taught, may be the crucial variable.

Motivational data were collected and are presented in the report. The most interesting finding was that when interviewed, the special class student who did well in the course tended to view school, his future job and his peers positively. That is, the positive correlations were evident with the posttest, not the pretest scores, unlike the regular class child.

The special class child's perception of his society, his peers and his school efforts may be more accurately related to his performance following a period in which he can demonstrate this ability than as he looks phenotypically in the usual academic classroom. The able child (in a learning potential sense) has strong hopes and aspirations which are grounded in a willingness to work in school, make realistic judgements about his future, and be seen positively by his peers. The less able child appears less adequate in his interview verbalizations.

The conclusion is inescapable. The hypothesis that high scorer and gainer special class children can be considered educationally rather than mentally retarded is borne out by the similarity of their learning pattern to that displayed by the dull and average IQ children who have done poorly in school. Though the special class children do know less about electricity to start with (or display more initial discomfort with, and avoid the evaluation task on the pretest), increased familiarity with the materials and the evaluation procedure, and exposure to the curriculum unit obscures these initial differences. These findings support the hypothesis that the high scorer and gainer special class child can learn satisfactorily with presentations appropriate to their ability structure. The evidence from this study of classroom learning suggests that their rate of learning is at least equal to that displayed by the low achieving regular grade agemates, at least in these special types of nonverbal learning situations. Whether this ability which is displayed is specific to nonverbal tasks, or can be tapped early in the child's school career and result in better progress in the verbal areas is a question that requires considerable study.

The data from this study suggests strongly that just as low achieving regular class children with dull to average IQs may require novel presentations and interventions by which to learn that will tend to minimize the negative effects of their poor literacy skills, so too would high scorer and gainer students with IQs in the educable mentally retarded ranges seem to be able to profit equally well from similar types of specially designed classroom interventions.

More importantly, for present practices and concepts of who is retarded, a learning potential strategy, used to discern the range of ability within a presumably homogeneous group of educable mentally retarded students, dictates a change in thinking about educational prognosis for the severe school nonlearners who demonstrate potential for learning. From this change, should emerge a drastically altered conception of how these children might be educated.

I. AN OVERVIEW and PLAN for the STUDY.

Thinking about the educable mentally retarded (EMR) of idiopathic or cultual-familial etiology has been beset by the curious paradox that the children so labelled during school do not appear as reported statistics once they reach adulthood. That is, the prevalence of EMRs falls precipitously once the period of compulsory school attendance ends. Follow-up studies of adults earlier labelled EMR have fairly consistently reported that these adults make satisfactory adjustments as a group when compared with control populations from the same socio-economic backgrounds. This curious paradox has given rise to an entire literature on "pseudo-feeblemindedness" which attempts to account post hoc for this discrepant set of findings. The child who does not do well in school is labelled, by implication, as a person who will have difficulty attaining economic and social self-sufficiency. But on the follow-up he is found to be economically and socially independent to the same degree as other adults who followed a similar underachieving route through school and came from the same kinds of backgrounds but were not labelled as EMRs on psychometric grounds. What is most interesting at present, with all the concern directed at the culturally disadvantaged, is that most EMRs come from similar socio-economic backgrounds as the culturally disadvantaged child. But, usually, due to a low psychometric score and persisting school failure (depending on the standards of the school system), they are classified as EMRs and segregated into special classes, though their potential for life success seemingly is greater than the EMR label would indicate. In the special classes, they are frequently taught materials that they have persistently failed to achieve success with. That they are poor school risks in the usual schoolroom is probably clear from the psychometric score (e.g., the Stanford Binet), but their later ability to achieve as satisfactory a place in the occupational structure as control non-EMR peers, suggests that the predictive criterion of psychometric scores such as the Stanford Binet is faulty and has yielded far too many false positive predictions.

This investigator has been engaged for a period of years in studying an alternate procedure for the evaluation of the "EMR" using a measure that may describe ability that is often untapped in the usual academic situation. A reasoning task (Kohs Block Designs) was selected to determine potential for learning when the task was sufficiently distant from a school type of task that the experience of failure attached to school would minimally inhibit performance on it. The child was given an opportunity to learn how to solve the task in an individual tutorial session following the pretest session. Degree of improvement in performance on the task was hypothesized to tap an intra-educable range of ability which might be related to the observed post-school differences in occupational attainment among EMRs. When the pretest scores were compared to the posttest scores, three patterns of response were evident. Some <u>Ss</u> had markedly improved their scores following coaching (<u>gainers</u>), others had not (<u>nongainers</u>). A third group performed exceedingly well on the pretest, contrary to expectations based on the IQ, and are referred to as <u>high scorers</u>. Both gainers and high scorers tended to solve problems

at similar degrees of difficulty, the difference being that the latter required no prior tutorial experience to solve the more difficult problems.

These different levels of response are evident consistently on other psychometric and learning tasks, strongly suggesting that the ability displayed on the block designs in this special assessment procedure was not task specific. Small differences between learning potential (LP) status groups are evident on verbally biased individual intelligence tests, e.g., the Stanford Binet and Wechsler verbal scale (Budoff, 1968b). The startling contrasts among special class children grouped within this trichotomy occur on tasks in which competence does not depend on proficiency in the verbal-conceptual or reading areas. High scorers and gainers perform significantly higher than nongainers on such tasks as Raven's Progressive Matrices, and Wechsler performance scale, attaining scores in the dull-normal to average ability range. The scores of the nongainers tend to be in the retarded ranges, and are not significantly different on these tests than what their verbally biased IQ scores would predict. High scorers and gainers also learn more rapidly and efficiently on a double alternation problem and a paired associate learning task (Budoff, 1967) and gainers tend to be less rigid than the nongainers and mental age controls, (Budoff and Pagell, 1968). Motivational data suggests that high scorers and gainers express feelings about themselves that are commonly described for school underachievers (Budoff, 1965; Harrison and Budoff, 1968).

The finding that large proportions of the psychometrically defined EMR group are able to reason adequately, as measured by the learning potential task, suggests that the high scorer and gainer "EMR" may be educationally rather than mentally retarded and may be extreme versions of the educationally disadvantaged child. They can succeed on reasoning and learning tasks and, perhaps, also in school, when the task requirements are congruent with their nonverbal manipulatively oriented skills, or, when they can perceive the tasks as nonacademically related, and work on them in a supportive, success-oriented context. Nongainers, on the other hand, do not exhibit this latent ability under these special conditions, and may be "truly" mentally retarded children, though this conclusion requires much further study.

The present grant (#32-31-0000-6019) was funded to test the educational implications of the learning potential argument as generated from the fairly consistent findings presented above. The study sought to test whether there would be differences in learning among special class EMRs on a curriculum taught in a laboratory science format which emphasized learning by manipulation of materials, and minimized the need for adequate skills in reading, as well as formal explanation and terminology.

The clear implication of the learning potential studies that the high scorer and gainer EMRs are educationally rather than mentally retarded suggests that the most suitable contrast group for them would be the educationally retarded (low achieving) child in regular class

who scores in the dull normal or average IQ ranges and lives in poor socio-economic circumstances similar to the special class student. Both the low achieving regular class child and the special class child have experienced prolonged periods of difficulty in learning in school, which serves to increase their avoidant response to academic subjects. This is a natural response for the failing child, especially when being successful in school seems to be so difficult for the slum child to accomplish. It is this pattern of avoidant behavior, with the reduced rate of acquisition of school related skills, e.g., in reading, and general knowledge, that is currently the focus of the innovative revolution in education. The special class child however, bears additional burdens. He has been stigmatized by being segregated into a special class, which he perceives, accurately, to be a terminal placement. In addition, he has more severe academic handicaps, and is regarded as incapable by his teachers and his peers. He probably has a more negative attitude toward school achievement, deprecates his own capabilities, and adopts a more severe nonlearning stance because of these more adverse motivational circumstances. It was assumed that if the proposed curriculum intervention could be made highly stimulating and motivating, this negative stance toward learning among the more able special class children might be undercut. The result might be a similar pattern of performance among high scorer and gainer "EMR"s and low achieving regular class students.*

The major hypotheses of the study were:

1. that high scorers and gainers from disadvantaged backgrounds would perform better than their nongainer peers from similar environments following teaching.

2. If the performance of the psychometrically defined EMR was compared with two groups of chronological aged (CA) contrast groups from similar social backgrounds, namely, adequate and low school achievers, the EMR high scorers and gainers would perform more poorly than the CA adequate achievers but as well as their low achieving CA peers. Nongainer EMRs were expected to perform more poorly than all the other groups.

These hypotheses attempted, then, to state summarily the educational implications of the prior findings with psychometric, reasoning and learning tasks within the "EMR" population. By hypothesizing underlying ability among high scorers and gainers EMRs and designing a curriculum that seeks to utilize these nonverbal reasoning abilities, the study seeks to test the most important implication of this research program, namely, that the high learning potential EMR, i.e., the high scorer and gainer, are educationally rather than mentally retarded, and they might more properly be considered "pseudo-feebleminded". If the null

*It was assumed that all the low achieving regular class students would fall in the high scorer or gainer categories, but this was not the case as will be evident below. Posthoc, then, this hypothesis is concerned with differences as a function of learning potential status hypothesizing no differences as a function of class assignment.

hypothesis were demonstrated, i.e., no differences between the high learning potential "EMR" and low achieving control, it would indicate that the high learning potential EMRs should not be placed in special classes, which tend to be terminal placements, because they are "dumb" or incapable in school. Rather it would indicate that individual programs of education need to be formulated which would aid these severe school failures, who are high school risks, to maximize their abilities in whatever course seems best for them.

The "pseudo feebleminded" controversy may be "solved" then, by the application of a learning potential assessment strategy, especially when new avenues to teaching these high learning potential "EMRs" are explored which will help them to maximize their potential strengths and strengthen their scholastic weaknesses. The argument underlying this research program suggests further that these kinds of revisions in curriculum formulation and presentation may also be applicable to under- or low achieving culturally disadvantaged children. On the most practical level, it was hoped that a concrete demonstration that the learning potential distinctions are educationally relevant might result in re-conceptualizing the educability of a significant proportion of the psychometrically defined EMRs. The low Binet IQ score clearly indicates that this child is a poor scholastic risk. The intent of this research project was to determine whether he is indeed relatively incapable of learning even in a specially designed educational setting, as the diagnosis usually implies to clinicians and teachers.

As a by-product of the study, data would be available on the utility of a learning potential assessment procedure for predictive performance in adequate and low achieving regular class children.

Educators in the past few years have been concerned with emphasizing the strengths of the educationally disadvantaged child in the classroom; among these strengths, the need to learn through concrete experiences and actual physical manipulation of objects ranks high (Reissman, 1962, 1965; Malkin, 1964). They take schools to task for stressing formal verbal skills -- such as reading and formal recitation in the class (Olsen, 1964) at the expense of concrete nonverbal learning and "informal" language skills the child is capable of, such as listening, observing, and speaking in a more casual environment. It appears that educators of the culturally deprived have the same argument with the schools that Budoff has-- the curriculum does not consider the individual child's optimal learning styles, motivations and the practical nature of his investment in education.

The plan of this project, which sought to test the educational implications of the learning potential argument, was to select a curriculum unit from the field of science in order to fulfill the requirements of a manipulative and a highly motivating unit. Some educators of the disadvantaged have singled out science as a natural subject for these students; the emphasis on first-hand laboratory experience promoted by modern science curriculum developers admirably suits the motor-oriented

style of the child, albeit with some necessary modification (Giddings, 1966). Likewise, the subject matter of science (again, if properly presented) intrigues disadvantaged children because of its natural relation to daily life and the power accorded science as an organizing force in life. Giddings summarized Reissman's (1962) argument that a deprived child's respect for education may emerge from science itself.

> This respect can perhaps be attributed to a physical, non-symbolic approach to life that has its beginnings in the disadvantaged child's experience; to a belief that science leads to an understanding of the complex world and a degree of control over it; to a belief that science is more closely related to ordinary activities than other academic disciplines; and to the child's ability to identify a career in science more readily than other careers. (Giddings, 1966, p. 438).

To test the hypotheses of the present study adequately, a satisfactory evaluation plan was required which could provide empirical evidence that different levels of comprehension were attained by the various learning potential groups, and to compare these levels with those of the control groups. Section II of this report reviews the arguments related to systematic evaluation of an educational curriculum unit. The special problems in the evaluation of barely literate adolescents are cited, and the evaluation strategy and the instrument developed for the present project are described.

The initial evaluation plan was developed for the trial teaching of one such science unit, which took place during the winter and spring of 1966 in a special school for mentally retarded students in Waltham, Massachusetts. The results of this trial teaching experience are presented in detail in Appendix A. The experience in evaluating this unit taught the project staff much which was useful in developing the final evaluation instrument used in the present studies, as well as basic principles in teaching slow learners, but the unit did not interest these children and presented difficulties in evaluation. Another science unit, Batteries and Bulbs,* which teaches simple concepts of electricity by having the student make bulbs light in increasingly complex circuits, was substituted. The subject matter was felt to be intrinsically more interesting to adolescents; the basic problems to be learned in the unit permitted a multiple choice evaluation scheme focussed around the basic problem "will the bulb in this set-up light?" The evaluation instrument formulated for this unit is described in Section II C.

Section III of this report describes the curriculum unit finally adopted for the study, the characteristics of the samples included, and

---

*These units were developed by the Elementary Science Study of the Education Development Center, Newton, Mass., and are distributed by McGraw Hill.

the procedures followed in presenting the curriculum unit, and in the data collections.

Section IV describes the results of the experiment, and the relations between the performances of the special and regular class students in their classrooms, on other psychometric abilities, and some interview data. Section V discusses the implication of the results for the learning potential hypothesis, and for educational practices with special class students.

Section VI describes a pilot study growing out of the results of the findings of the major study, namely an attempt to determine whether or not the student's actual physical manipulation of the materials in class was the critical factor in facilitating the comprehension displayed in the unit. It also relates the results of the pilot study to those of the main study in an attempt to establish the reliability of the students' performances on the electricity unit.

Section VII summarizes the conclusions and implications of the entire study.

II. PROBLEMS INVOLVED IN EVALUATING STUDENT PROGRESS AND THE PROJECT EVALUATION SCHEME.

A. General Controversy re: Formal Evaluation

The subject of objective evaluation in science education is a controversial one today. At one extreme are the proponents of non-formal evaluation who contend that the objective charting of short term behavioral changes due to a science unit is insignificant; what is really important, the long-range objectives, cannot be measured in this way (Atkin, 1963). They also believe that the firm definition of unit objectives, prerequisite for formal evaluation, dampens the creativity of the curriculum developers, and that many positive outcomes of science units (particularly those developedin the past ten years) such as "interest in science" and "willingness to experiment" are overlooked in objective testing because their manifestation is so subtle.

Elementary Science Study, the source of the science units used in the present project, is among the strongest apostles of the anti-formal evaluation school. The group uses methods such as informal observation and teachers' comments to judge the worth of their units, but holds formal evaluation worthless. "We are not involved in evaluating these materials through objective testing, our belief being that such a procedure is limited by the validity of the evaluative instrument used." (Response to section on "Tests Being Developed" in <u>Third Report of the Information Clearinghouse on New Science and Mathematics Curricula</u>, 1965, p. 31).

A more moderate position in the controversy is maintained by those teachers and educators who look for alternative ways to evaluate but still consider objective testing valid and valuable. J. Myron Atkin criticizes the limitations of objective evaluation and uses teachers' and scientists comments in assessing the University of Illinois Elementary School Science Project in astronomy, yet, he includes comprehension tests in his evaluation as well (Atkin,1963). William B. Reiner calls for a needed review of a wider spectrum of evaluation techniques - laboratory tests, self-report inventories, informal observations, etc., but he also seeks the production of better objective measures which will test higher levels of comprehension and be better adapted to children's reading skills (Reiner, 1966).

At the other end of the scale, are those who uphold the validity of traditional objective evaluation of measurable outcomes and who continue to develop appropriate tests. Persons holding this view seem to divide themselves into two groups. One group is interested in making better objective tests for more critical levels of scientific understanding than are usually represented in science tasks. Often aiming towards improvement of tests of public school science, these individualsdecry the widespread process of testing mere recall of content. They generally base their ideal tests on the classification suggested in Benjamin Bloom's <u>Taxonomy of Education Objectives, Handbook I, Cognitive Domain</u>,and thus divide knowledge into six different categories: 1.) Simple knowledge, usually factual knowledge; 2.) Comprehension, or the ability to translate, interpret, and extrapolate knowledge learned: 3.) Application -

7 .

the "ability to handle problem situations by relating the appropriate generalizations or abstractions to information which has been presented or recalled", (McFall, 1964, p. 104); 4.) Analysis, the breakdown of material into its constituent parts and the detection of the relationship of these parts; 5.) Synthesis, the converse of analysis, the putting together of an idea or of facts, and finally; 6.) Evaluation - the making of judgements by means of internal or external criteria (see Bloom, 1956; also Hedges, 1966; McFall, 1964; Nelson, 1959; Lombard, 1965). It is obvious that certain of these categories of learning, particularly the last four, can be taught only in higher level science courses in high school and beyond.

Evaluators in the second group are developing objective tests for measuring acquisition of processes of science, usually among children in the lower grades. Elizabeth Hagen, in collaboration with Science Curriculum Improvement Study (SCIS) curriculum constructors has recently developed a series of objectively scored process tasks for the project's first unit, Material Objects (Hagen, 1966). The American Association for the Advancement of Science (AAAS), is developing a K-12 curriculum based exclusively around processes. In this curriculum classes are lab-oriented -- children learn the process of classifying through actual categorizing of different objects -- and teacher's appraisals, one form of the AAAS evaluation, are similarly focussed on active work with the materials. For example, to appraise a lesson on observing magnets and magnetism, children are given, among other things, two magnets and a number of paper clips and asked to determine which magnet is the strongest. Evaluators also tap progress of children throughout the year on the defined processes through tasks related to their classroom work; children are rated by a "check list of competencies".

AAAS evaluators strongly uphold the validity of the objective tests. One of their number, Henry Walbesser (1963), deprecates the anti-formal evaluation school with their indistinct objectives:

> There exists a class of curriculum designers who do not wish to be (as they often contend) "pinned down to specific behavioral changes". As a direct consequence of this lack of commitment to a behavioral description, they frequently do not specify the performances they expect the learner to accomplish. Even in those rare cases where specification is made, the objectives are all too often couched in terms of educationally pleasant but meaningless language such as "the child will appreciate" or "the child will be able to deal more effectively with" or "the child will understand better". It is impossible to construct direct measures from such linguistic drivel (p. 297)

In the same discussion, Mr. Walbesser also criticizes their requirements for creativity, ultimate flexibility, and freedom from rigorously defined objectives, and explains the rejection of such a position in AAAS evaluation:

It has been contended that placing such severe constraints upon the system of experimental curriculum construction may tend to inhibit the "creativity" of those involved in writing the materials. However, since the proponents of this thesis have as yet presented no experimental evidence to support their claim, it has not affected this strategy for curriculum evaluation (p.299).

In choosing an evaluation strategy, the research team considered the view of the anti-objective testing school of thought, but in reviewing the hypotheses, it was decided that the demonstration of heterogeneity of ability within the special class students and between special and regular class students could be definitively made only through objective assessment in a formal pre-posttest design. However, it was agreed that such a procedure might not adequately describe the behaviors and motivations which aided or hindered the science learning. A classroom observation schedule requiring formal ratings and anecdotal descriptions was added to the evaluation scheme.

B. Problems Associated with Evaluation of Curricula with Partially Literate Adolescents.

In order for the testee to give his best possible performance, he must be able to fully understand the testing procedure. Tests for low achieving and special class students from working class or disadvantaged backgrounds must minimize the verbal and reading aspects, as these are commonly areas of poor skill competence. Psychologists and sociologists have long commented on the unfair bias of the verbal sections of intelligence tests, and the fallacy of judging children from disadvantaged and other subcultural groups by using terms familiar mainly to middle class children. If the child cannot comprehend the word, how can he define it or use it in an analogy? Davis and Eells (1956), Cattell (1940), and others have attempted to develop culture-fair or culture-free intelligence tests. Another solution has been to eliminate verbal testing entirely, and depend on performance tasks to assess ability. Budoff's learning potential procedure in which children are given a performance task, coached in one style of solving the problems, and then retested to assess ability to profit from the learning experience (a general definition of intelligence) is a novel instance of this latter approach.

Though several educators realize the disadvantaged child's difficulty with reading tests, they have been slow to apply this fact to their achievement evaluations. The field of science would seem a natural place for the production of nonverbal tests for culturally different children, but such development has not taken place. The few nonverbal science tests that have been developed have been laboratory tests for average and above average ability high school students (see Reiner's references, 1966), or for primary school children -- natural "non-readers". As has already been mentioned, AAAS and SCIS, in a less extensive way, use laboratory situations to judge young children's prowess in processes - classification, observation of differences, etc. In both cases, children are

asked to perform certain tasks and are then graded according to predetermined criteria. Boener (1966) has developed a nonverbal multiple choice test for judging children's science concepts in the kindergarten, first and second grades. She makes use of pictures and diagrams, and uses verbal instructions to facilitate the children's choices. Reiner (1966) makes a bid for nonverbal testing techniques and reviews good methods for such tests:

> The problem of the tests can be presented by using models, charts, posters, diagrams, and toys. By having the pupils respond to spoken instructions rather than written ones, the reading burden can be virtually eliminated. Answer sheets can be prepared to conform to the questions. The pupil responses can be limited to marking an X or circling a number or letter. The challenge lies in inventing good problem situations (p. 336).

Unfortunately his remarks are directed toward tests for primary school children and not towards older and barely literate students who may also need a similar evaluation approach. There are few precedents for nonverbal achievement tests for this group.

C. An Initial Evaluation Plan for the Present Study

The initial evaluation strategy for the science unit was a laboratory test in which children carried out tasks and commented on what they had done. It was a good test for partially literate students in that it was basically a series of performance tasks, and the children could readily understand what was expected of them. However, it provided too much opportunity for learning the topics covered in the curriculum unit on the test itself, and did not minimize the role of verbal expression enough - facts which clouded the assessment of what the children had learned as a result of the classes they attended.

The laboratory format was suggested by the obvious measurable outcomes of the science unit. The first unit selected by the project staff was <u>Kitchen Physics</u>, (1965), a unit created by the Elementary Science Study (ESS) of the Education Development Center (Newton, Massachusetts).

Initially, the attempt was made to review the unit, its objectives, content, etc., so that an evaluation plan would be oriented toward the kind of information or processes the pupils should have mastered by the end of the unit. A close examination of <u>Kitchen Physics</u> revealed that the major content concepts were tension-adhesion and viscosity. It was noted, however, that these were in reality very poorly delineated; they could both be summed up in the basic theme of the unit - water grabs together better than soapy water. The other concepts of the various weights and absorption and evaporation rates of liquids were better defined, but the whole unit did not hang together thematically.

Processes of science, on the other hand, were well developed in <u>Kitchen Physics</u>. Observation skill, prediction from previous experience,

translation of a problem into numbers, the notion of error, and the use of equipment were all integral parts of the unit. After consultation with ESS, a set of testable categories were formulated: 1. skill of Observation, 2. ability to offer a Solution or reason behind observed phenomena, 3. Prediction on the basis of previous knowledge gained in the test, 4. Reason for the prediction given, 5. use of Equipment, 6. understanding and use of a Balance, 7. understanding of the equivalence of Weights (related to 6), and 8. knowledge of Experimental Error. Although these categories emphasize processes, several of the areas (particularly 2, 3, and 4) depend upon knowledge of principles as well.

The emphasis on process skills among the objectives of Kitchen Physics suggested a series of laboratory tasks as an evaluation instrument. Such a test, administered individually, would enable the children to demonstrate their process skills and knowledge in an active manner and thus maximize their response. A typical sequence of the test as finally devised is given below. The juxtaposition of content and skill areas can clearly be seen. (Phrases in parentheses indicate the category of response.)

The student was first asked to place two drops of water on a glass plate with an eye-dropper (Equipment), push them together and describe what happened (Observation). He then repeated the performance with soapy water, stated any difference observed between the water and the soapy water (Observation), and hypothesized what would happen with cooking oil in the same series of tasks (Prediction); he was also asked to offer a reason or solution for any differences observed between the water and the soapy water (Solution). Next he was presented with a transfer task to which he might apply the information already learned above. To a level medicine cup of water he was asked to add "all the water you can with the eye-dropper until the first drop spills" (Equipment); he then did the same with a level medicine cup of soapy water, first hypothesizing whether it would heap to the same level, higher, or not as high as the water (Prediction). Following this task he was asked to observe the two levels, detect the higher one and offer a solution as to why it was the higher (Observation; Solution). (See Appendix A, for a full description of the test instrument).

The evaluation scheme fulfilled many of the requirements for a good test for partially literate students. Primarily, it eliminated the necessity for reading and the punitive aspects of a pencil and paper test. The requirements were easily understood by the children and many of the tasks engaged their attention while allowing them to perform without any verbal handicap.

On the other hand, the scheme was felt to be inadequate for several reasons. Many difficulties centered around the fact that in a laboratory procedure, children have to be shown how to work with problems and equipment. Much thunder and novelty was stolen from the classes as the children experienced much of the course, in nucleo, during the pretest. Again, the laboratory activity enabled some children to learn more than others during the pretest and thus start out with a higher score than

their beginning knowledge warranted. Other problems were posed by the fact that although many of the test items (in the categories of Equipment, Observation, Balancing) were purely performance items, many others (in the categories of Solution and Reason for Prediction) depended on the children's ability to verbalize their experience. An objective marking system was set up to score these verbal responses, but it was difficult to apply the categories to the actual responses. Typically, the verbal answers were not clearly and specifically enunciated, and sometimes tended, when they were incorrect, to contradict the correct understanding shown in the child's performance on the test.

In reviewing the results of the Kitchen Physics study it was decided that the laboratory test, however attractive in certain ways, would have to be abandoned. Another measure was needed which would still offer direct experience with the material but which would bypass any need for verbalization and lend itself better to objective scoring. A short answer multiple choice test was the form proposed.

For several reasons, preeminently the children's lack of interest in the subject matter and these problems in evaluation, Kitchen Physics was discontinued. Its replacement, Batteries and Bulbs (1966) was chosen because it was felt to be intrinsically exciting to the adolescent subjects, was rich in material, concepts and in specific objectives, and because its basic teaching questions "will the bulb light"?, and "how bright is the bulb"? provided a natural format for a short answer objective evaluation.

D. The Evaluation Plan Adopted for this Study.

The evaluation scheme developed for Batteries and Bulbs solved the problem of the children's verbal and reading deficiencies, while giving them a modified first-hand exposure to the materials they were questioned on. The test, individually administered in two parts, consisted of questions on a number of electrical set-ups. The format was simple; the set-ups were presented as multiple choice problems ranging from easy to very difficult. In the first half of the test (Boards) the set-ups were made of actual deadened batteries and bulbs mounted on masonite peg-boards. It was felt that this visual exposure to the materials of the course maximized a child's chances of demonstrating comprehension of the concepts of electricity. The second half of the test consisted of realistic and schematic diagrams which both tested the child's knowledge of electrical diagramming and provided a differential measure of his ability to abstract from the concrete materials. Figure I presents pictures of representative boards as they appeared to the student. Appendix B is the booklet used in the Diagrams section of the test.

In subtest I of the boards (Simple Circuits) Ss were presented with twelve multiple choice items, each consisting of four set-ups of different combinations of one battery, one bulb, and two wires. Each set-up was labelled by letter; the student was required to pick out the one

set-up which would work (the complete circuit) in each group. S was told that "All batteries here are dead, and the wires and bulbs pasted on. If the batteries were working, one bulb in each group would light. Which do you think it would be - A, B, C, or D" ? On three of the twelve problems, they were asked to give a reason for their choice - an attempt to get an additional demonstration of comprehension, (See Figure 1A).

The second sub-section of the boards (Complex Circuits I) covered a variety of electrical topics - parallel and series battery circuits, parallel and series bulb circuits, solid and liquid pathways, etc. A demonstration of S's knowledge was again gleaned from a simple multiple choice type of question. S was introduced to a "standard circuit" - one battery connected to one bulb (which was lit); he was shown how "standard brightness" could be made with or without a bulb-holder. S was then shown the various single set-ups (a total of 36 items) on the masonite pegboards, in which particular bulbs were marked "A", and asked, "In each set-up, will the bulb marked "A" be the same brightness as standard brightness, brighter than standard brightness, dimmer (or darker), or not light at all"? Directions were repeated from time to time, especially when S showed a tendency to fixate on an answer (See Figure 1B).

Problems in the third section of the boards (Complex Circuits II) were drawn from the same topic areas as those of the second section, but the format was slightly different. There were two set-ups within each each question and children were asked to compare the brightness of a bulb in the first set-up with one in the second. "If the batteries worked, will bulb "A" (in the first set-up) be brighter than bulb "B" (in the second), will both be the same brightness, or will neither one light"? To solve the problems, children had to identify the distinguishing feature or features of each set-up, some of which were fairly complex, and then put all the facts together to make the comparison. There were twelve items in this section, (See Figure 1C). Four "why"? questions were asked which required a verbal response in this section.

The second half of the test was made up exclusively of diagrams - two sections of pictorial representations in the form of actual batteries, bulbs, and wires, and one section of schematic electrical diagrams. It was presented in a second session, individually. Ss were given test booklets and pencils for the diagrams portions of the test (See Appendix B). Section I was an exact replica of the first section of the boards; S was asked this time to place an X under the set-up in each group that would light. Section II of the diagrams duplicated Section II of the boards with the elimination of several problems on different kinds of bulbs and wires that would require too complicated a symbolism. There were 28 items in all. S was presented with the standard brightness circuit (a working circuit consisting of one bulb and one battery) and asked to compare the relative brightness of bulbs in the pictured set-ups; he indicated his choices by marking "S" for "same" (or standard), "B" for brighter, "D" for dimmer, and "N" for not light at all, in the appropriate blank. Any abstract representations, like that used for a bulbholder, were explained to him. Part III diagrams were schematic abstracts of Part II diagrams (with a few examples schematized from Part III of the boards).

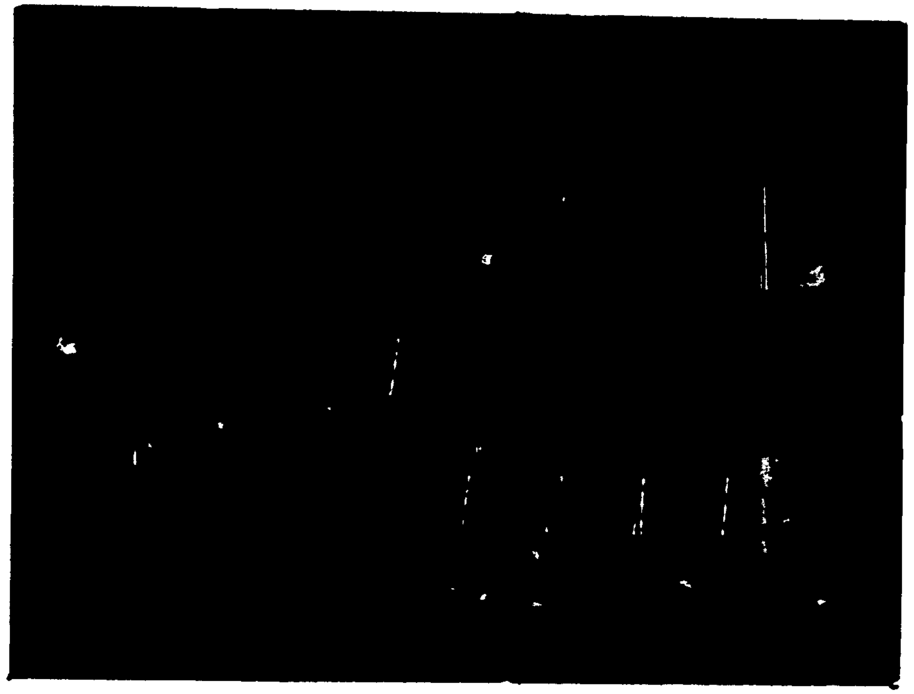


Figure 1A. Boards Section, Simple Circuits

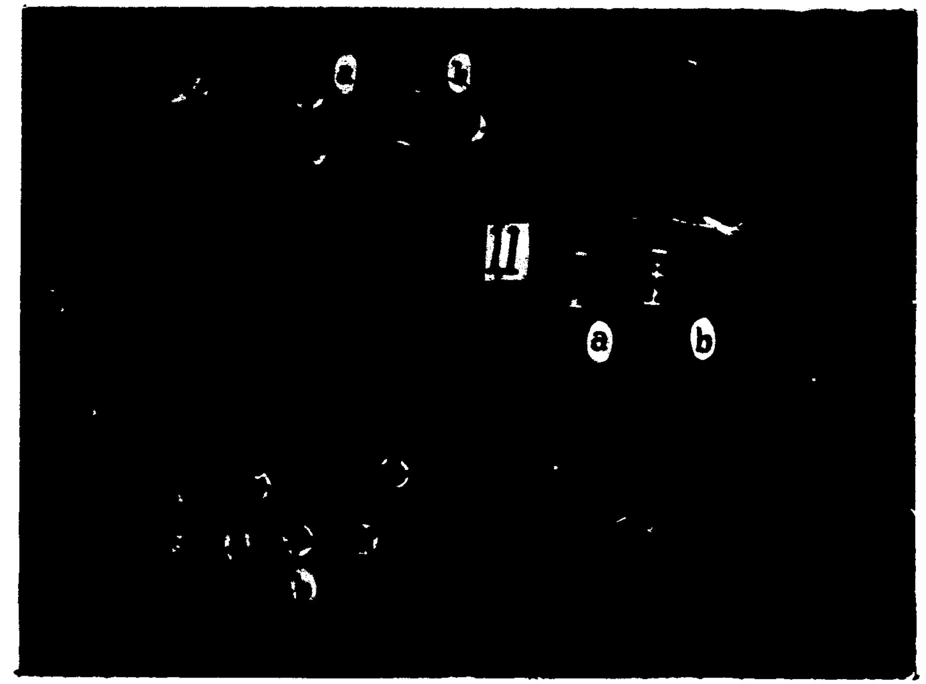


Figure 1B. Boards Section, Complex Circuits

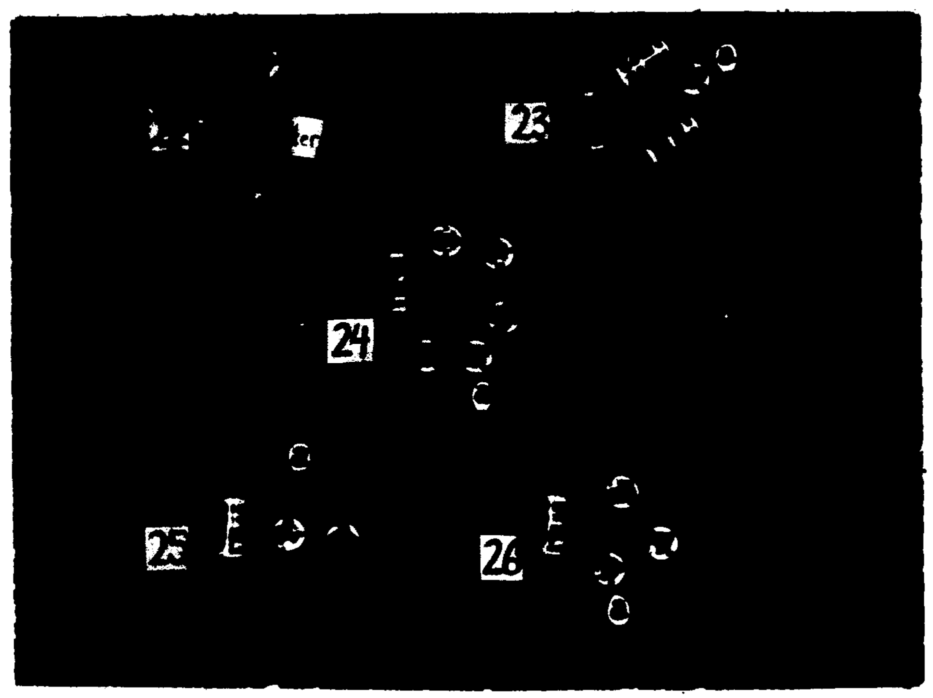


Figure 1C. Boards Section, Complex Circuits 2.

There were twenty examples in all. Translation of the realistic into schematic diagrams was carefully explained to S; he was required to identify a simple circuit drawn in schematic format, consisting of a battery, bulb, and wire. As in Diagrams section II, brightness comparisons of the bulbs in the schematic diagrams were made with "standard brightness", children indicating their answers with "S", "B", "D", or "N".

The project staff felt that the final test maximized the children's chances of displaying their knowledge of electrical concepts and applications. Problems of both reading and verbal expression eliminated any need to read. In addition, the display of the actual materials employed in the unit on the masonite peg-boards provided the best exposure to the problem situation possible without having the children actually manipulate the equipment. The student was given the opportunity to demonstrate his knowledge without having to translate it into any abstraction, diagrammatic or verbal. Alongside this concrete task, however, the two diagram sections, one containing realistic diagrams and one schematic diagrams, were included. These sections were intended to discern whether the students could solve problems in the three-dimensional format better than in the equivalent pictorially presented diagrams on the pre and posttests, and also to see whether they could effectively translate from these formats to the schematic diagrams. Finally, although several verbal "why?" questions were asked, they were not included in the total test score or in grading S's response to the specific item. Rather, these open-ended questions were to be scored separately and constituted a means of sampling whether the S could also correctly verbalize his behaviorally demonstrated understanding.

III. THE CURRICULUM AND PROCEDURES EMPLOYED IN THE STUDY

A. Description of the Curricula Unit Used

1. Introduction to selection of the unit used.

Because of the superabundance of science materials that have been produced in recent years, the research staff had a wide field from which to select the appropriate unit. Since the late fifties, such groups as the Science Curriculum Improvement Study (SCIS), the American Association for the Advancement of Science (AAAS), the Illinois Elementary Science project, and the Elementary Science Study (ESS) of the Education Development Center (formerly called Educational Services, Inc.) have been engaged in production of improved science units for elementary and junior high school age children. Groups involved in new approaches to high school science include the Chemical Bond Approach (CBA), the Physical Sciences Study Committee (ISSC), and the Biological Sciences Curriculum Study (BSCS). The new projects place differing emphasis on the content to be learned or the scientific know-how, or inquiry attending it. All, however, share an emphasis on direct experience in the laboratory, doing science rather than memorizing the achievements of other scientists - a basically inductive approach to learning emphasizing the student's own manipulation of the materials.

It was decided that a cirriculum developed by the Elementary Science Study (ESS) would be most apt for this project since their units are directed towards elementray and junior high school students. Their threefold purpose is:

> . . . to contribute to a more balanced curriculum by bringing science into the classroom of the early grades; to arouse the curiosity of all children, and at the same time to cultivate their desire and capacity for inquiry; and to supply teachers with a variety of carefully thought-out and tested materials that they can use to build the elementary science curriculum best suited to the particular needs of their pupils (Goodlad, 1966, p. 51).

ESS's many units, Growing Seeds, Gases and Airs, and Kitchen Physics, cover both content of a particular area and scientific processes; their approach is fiercely inductive; children learn by concrete manipulation of materials. Free exploration of ideas springing from the "lab" work is central -- "Materials selected for study are those which inherently allow for a flow of ideas originating from the curiosity of children" (Third Report of the Information Clearinghouse on New Science and Mathematics Curricula, 1965); formal explanation or terminology is rejected and children

are instead urged to explain their findings in their own words. It is obvious how such a philosophy lends itself to the needs of this particular project.

In the winter of 1966 when the present grant was funded, Kitchen Physics seemed the most promising of the ESS units. By experimenting with water, soapy water and oil, heaping, dropping, "beading", and weighing them, children are introduced to such properties of liquids as viscosity, weight, and surface tension. They also learn processes of science - observing, predicting, concluding, using equipment, etc. A close analysis revealed that these processes are actually better defined (and more measurable) than the content areas. Since the unit has been used successfully with children from the fifth to eighth grades, it seemed appropriate for our experimental and control group of young adolescents. Finally, subject matter is "drawn from a child's own environment" (Teacher's Manual, Kitchen Physics, 1965, p.1) and was felt to be particularly appealing to educationally disadvantaged children.

A trial teaching run of Kitchen Physics was conducted in the spring of 1966 at the Royal E. Robbins School in Waltham, Massachusetts, which is composed entirely of special classes. The results of that trial teaching, a detailed description of the evaluation instrument formulated, preliminary empirical findings, and a critique of the Kitchen Physics unit, as taught and as evaluated, are presented in Appendix A.

In considering the unit for use in the studies projected, several serious defects were evident; these are elaborated in Appendix A. The most serious of these was the children's lack of interest. The older trial teaching class (mean age 15.2 years) was too sophisticated for the homemade, "scientific" materials of Kitchen Physics. Though the disadvantaged child respects science, (Riessman, 1962, p.13) he has certain well defined notions as to what science is; it may involve machines and scientific instruments, but not plastic bottles and wax paper. The younger group (mean age 13.3 years) found liquids better for messing than experimenting. Reviewing these facts and the evaluation problems already cited in the previous section, it was decided to switch to a content rather than process oriented unit which would be easier to evaluate, and more interesting and acceptable to these adolescents.

2. Description of curriculum unit adopted for the study

As a replacement for Kitchen Physics, the project staff found a newly published ESS unit, Batteries and Bulbs (1966). An electricity unit, Batteries and Bulbs again conveys content

through the method of inductive generalization and exploration.

> Batteries and Bulbs is an introduction to the study of electricity and magnetism in the course of which each child makes experiments with his own simple equipment (Flashlight batteries, small bulbs, various kinds of wire, compasses, magnets) and draws conclusions based on the results of his observation of these experiments. Children investigate such things as ways to light several bulbs with one battery, what happens when more than one battery is used, whether varying lengths and types of wire influence the brightness of bulbs.
>
> In the course of these investigations, the child gains the personal experience which makes the electrical behavior of matter an acceptable and familiar phenomenon. From the very beginning, using a battery and a piece of wire to light a bulb, experiments suggest questions which in turn suggest new experiments. Results are predicted. Students check experiments, compare and discuss results and propose explanations (Teacher's Manual, Batteries and Bulbs, 1966, p.1).

Within this framework, children learn what a simple circuit is and what requirements are necessary for making a bulb light. They study the results of varying amounts of voltage and amperage in an elementary way by making and observing different circuits with bulbs and batteries arranged in series or in parallel. They note the effects of resistance in wires of different materials, lengths, and thicknesses, and in different kinds of bulbs. Conductors and insulators, both solid and liquid, are observed, and lessons in diagramming with real "electrical language" are included in the unit.

Language is again kept simple -- children measure by a unit of one bulb connected to one battery called "standard brightness"; they learn few formal electrical terms, and formal explanation of concepts such as electrical resistance is minimized. No attempt is made to offer a comprehensive theory of electricity. Batteries and Bulbs is a course on applications of electricity -- what variations produce what effects in circuits -- rather than a theoretical introduction. Understanding is expected to develop as a result of working with the elements that compose an electrical circuit.

The project staff felt that electricity would be exciting to adolescents; that the unit was far richer in content than Kitchen Physics, and being richer, would have better measurable objectives. After deciding to adopt Batteries and Bulbs, a trial teaching run of the unit was held in the Robbins School in the autumn of 1966. During this run, two important developments took place:

a. A core unit was designed.

The course as presented in the Batteries and Bulbs manual is ectremely flexible and covers as little or as much as the individual teacher wishes to, or the class can manage. From the available topics, a twenty-six lesson Batteries and Bulbs unit was concretized for the actual experiment to be held in the winter-spring of 1966-67. The topics and the time allotment for each topic was approximately as follows:

Complete Circuits, Incomplete and Short Circuits	6
Parallel and Series Bulbs and Batteries Circuits	9
Resistance - Wires	2
Resistance - Different bulbs	4
Conductors	3
Miscellaneous (Review, etc.)	2

Appendix G presents the student's workbook for this unit which gives a good idea of the style of presentation and emphases of the curriculum unit.

b. An evaluation scheme was formulated; this has already been described in the previous section (Section II)

B. Design of the Experiment.

The study was designed as a pre-posttest comparison of knowledge of electricity as a function of the teaching intervention. Knowledge of electricity was determined by scores on the evaluation instrument and constituted the dependent measures. Four groups of students were included in the study: taught and nontaught special class students, and taught adequate and low achieving regular class students of the same CA. Each group included high scorers, gainers, and nongainers, and boys and girls. It was hypothesized that:

1. Following teaching, the taught special class students would have greater knowledge of electricity, i.e., higher evaluation test scores, than the nontaught special class students.
2. That high scorers and gainers would learn more from the course than nongainer special class students;
3. That the total of regular class students (adequate achievers) would learn more about electricity than taught special class students.
4. That special class high scorers and gainers would perform similarly to the low achieving regular class students following teaching;
5. It was not clear whether learning potential status -- i.e., high scorer, gainer, or nongainer -- would predict different levels of performance among regular class students as it does among special class students.

The original study, as planned, and funded, sought to investigate the effects of a productive thinking unit which was said to stimulate verbal fluency, on learning the science materials.

This unit was formulated by Sue T. Rouse (1965). It was hypothesized that if this unit could stimulate productive thinking of EMRs, there should be an effect on their learning in the science unit. This effect would be reflected in higher posttest scores on the various sections of the evaluation instrument, and perhaps, most particularly, in clearer, more directed responses to the verbal questions. This unit, as described by Rouse, was administered in a controlled study with EMR experimentals and controls, prior to the start of the electricity unit, in an effort to replicate Rouse's positive findings in South Carolina with Boston samples. There proved to be no differences in productive thinking scores between experimental and control groups on the subtests of Torrance's battery that were employed by Rouse in her study. As a result of these negative findings the decision was made to drop this set of questions from the major study described in this report. A report of this replication study is presented in Appendix C. It will appear shortly in the American Journal of Mental Deficiency (1968).

C. Method of Selection of Subjects and Description of Sample.

The students who participated in the study were drawn mostly from one junior high school in Boston that is largely composed of white Irish Americans. Additional EMR nontaught controls were included from a junior high school that is similar in composition to the school at which the classes were conducted. All the students in the special classes for the educable mentally retarded were included in the subject pool; the only exclusions were for demonstrable cerebral involvements or extremely evident emotional disturbance, i.e., having a suspicion of psychosis but not necessarily a severe behavior problem. The regular class children, who served as controls, were drawn from threeseventh grade sections: a low academic track section, and two low general track sections.* All the children in these sections who had failed two or more major subjects following the close of the first marking period (ending October 31, 1966) were screened as potential participants. It was supposed that some students allotted to the study with poor initial records would markedly improve their records by the close of their initial school year in junior high school providing samples of adequately and low achieving CA controls. This was indeed the case, as will be indicated below. The regular class students, then, provided a contrast group of good and poor students, in the 80-110 IQ range.

All the students, EMR and regular class students, were administered the learning potential procedure. This procedure involves individual administration of the sixteen test designs and five coaching designs three times: prior to coaching, one day,

*These were sections that had two study periods per week which were scheduled over two of the three days on which the project teacher was available.

and then one month following coaching. The usual instructions for administering the Kohs Block Design Test were used at each test session (Kohs, 1923). A sample problem was demonstrated by the examiner and the subject had to construct it correctly before the remainder of the items were presented. Testing was discontinued after three successive failures.

The test series of problems consisted of fifteen of the original seventeen Kohs block designs (Kohs, 1923), arranged in order of increasing difficulty, up to and including five designs with sixteen blocks. Designs 3 and 7, were omitted. Design 7 from the Wechsler Adult Intelligence Scale (WAIS) was added to make a total of sixteen test designs. The designs were printed on white cards double the scale of the original Kohs designs and equal in size to the completed design when composed of the usual one inch blocks. The design of four blocks was drawn as a two inch square, one of nine blocks as a three inch square. This modification was suggested by Goldstein's (1945) adaptation as a means of simplifying the task. The four colors, i.e., red, white, blue, and yellow were retained. The usual one inch cubes were used as blocks.

Five designs were used in coaching. They consisted of design 9 from the WAIS, items C and 5 from the WISC, and designs 3 and 7 from the Kohs series. The five coaching designs consisted of three four block designs and two nine block designs which were included in the test series but not in the test score total. The coaching designs were printed on white cards in the same dimensions as the test designs (for further details of this procedure, see Budoff and Friedman, 1964).

Based on the patterns of performance displayed on the learning potential task, the students were assigned a learning potential status. Students were considered gainers, if they met the criterion of solving at least four more designs on the post-coaching sessions than on the pretest, an increase in score equal to three times the mean increase of the noncoached control group (1.2 designs) (Budoff and Friedman, 1964; Budoff, 1967). High scorers successfully solved one of the difficult 9 or 16 block problems in the upper third of the test series prior to tuition and were considered to have demonstrated learning potential by performing better than would be expected on the basis of their IQ rating. Nongainers included all those coached Ss whose pre to posttest score change was less than four designs.

Table 1 presents the means and standard deviations for IQ, CA, and a measure of social class (rating of occupation of principal wage earner). It also presents a summary of two-way analyses of variances of these factors for the three groups, each with three levels of learning potential.

Table 1.

MEANS, SDs & F-RATIOS FOR EMR, TAUGHT AND NON-TAUGHT, AND CA CONTROL SAMPLES FOR IQ, CA AND OCCUPATIONAL RATING OF PRINCIPAL WAGE EARNER

	N	IQ $\bar{X}$	IQ SD	CA $\bar{X}$	CA SD	Mean Occupational Rating (Social Class Measure) $\bar{X}$	Mean Occupational Rating (Social Class Measure) SD
Taught EMR							
High Scorers	9	71.444	11.114	177.000	5.123	2.111	0.690
Gainers	12	73.000	5.291	170.833	10.435	2.167	0.756
Non Gainers	12	68.167	5.306	165.167	9.907	2.083	1.136
Non-Taught EMR							
High Scorers	6	76.667	3.502	170.167	8.886	2.111	1.304
Gainers	8	69.875	6.707	166.875	9.326	2.000	0.756
Non Gainers	12	63.583	9.802	173.583	11.000	2.250	1.302
Taught CA Controls							
High Scorers	17	94.235	11.405	158.94	12.11	2.353	1.320
Gainers	8	94.875	6.707	154.38	11.66	1.875	1.952
Non Gainers	8	85.625	8.634	160.25	11.47	1.375	1.506

	df	IQ	CA	Social Class
Learning Potential Status (LP)	2	21.04^{2}	.39	<1
Teaching Groups (G)	2	61.39^{2}	17.59^{2}	<1
LP x G	2	<1	2.16	<1
Residual Mean Square	83	71.77	107.13	1.32

[1] $p < .05$

[2] $p < .01$

Significant F-ratios were obtained for the learning potential groups on IQ ($p<.01$) regardless of whether they were in the special or regular class. Examination of the means indicates that high scorers had higher Binet-type IQs, gainers' IQs were higher than nongainers, but lower than those of the high scorers'. These differences are similar to those reported by Budoff (1968) after a study of large samples of special class children by learning potential status. The significant difference in IQ among the teaching groups ($p<.01$) was expected, and due to the higher IQs of the regular class students.

There were also significant differences in CA among the teaching groups, ($p<.01$), as the special class students averaged a year older than those in regular class. This CA imbalance resulted partially from the fact that no eighth or ninth grage classes were available in the school with two free periods to match with the older special class students, since the special classes included all the eligible students in the district between 12 and 16 years of age. Where possible, the older special class students were excluded from the study to reduce the CA gap. But this procedure was not sufficient to reduce the discrepancy since the pool of available students was restricted.

The occupations of the mothers and fathers of the children were obtained from the school records. There were no significant differences between any of the groups by social class when this was determined by rating the principal wage earner's occupation according to Turner's classification of occupations (Turner, 1964). It should be noted that the mean occupational ratings fall into the category Turner designates as "skilled laborer". With the exception of the regular class gainer group, the standard deviations indicate that all the children's families would fall into the laboring or lesser white collar categories (a maximal rating of 4 on a scale of 1 - 9).

D. Procedures.

Prior to the start of the classes, all the regular and special class students were administered the Batteries and Bulbs test as described in Section II D. The EMR nontaught controls at the school in which the classes were conducted were administered the pretest at the same time. The posttest was administered immediately following the cessation of the classes to all the taught and nontaught students at the first junior high school. The pretests on the Batteries and Bulbs evaluation unit were collected somewhat later for the EMR nontaught controls at the second junior high school, but these subjects were posttested somewhat later. A relatively similar time interval was maintained between the two tests for the taught and nontaught subjects. All the regular class students were included in the classes.

Thirty-five special class students received the experimental curriculum, and 17 EMRs from the same junior high school and 23 EMRs from the second junior high school were included as nontaught controls. For purposes of analysis, two of the taught EMRs were eliminated from the final sample due to excessive absence from class. Twenty-six Ss from the nontaught groups were matched to the remaining thirty-three taught special class students on the basis of IQ, CA, and pretest scores. Thirty-six children from regular class, received the curriculum; three of these were later eliminated because of excessive absence.

The special class children were assigned to one of three classes: an all boys, an all girls, or a mixed sex class, since the sections of regular class children available for use as controls were in these combinations. The taught EMRs were divided into three classes of 11 or 12 each; the taught controls into three classes of 11 to 13 each (division prior to elimination of high absentees in both groups). No attempt was made to balance all the classes exactly for composition by any factor but sex, because the various regular class sections had students that differed in IQ distribution, though they tended to be more homogeneous in CA. The EMR classes were balanced by distribution of learning potential status. The curriculum as described was kept fairly constant over the six classes.

The original grant proposal envisioned that the teaching would be done by the classroom teachers. However various considerations forced a change. The electricity unit, which was finally adopted required more specialized knowledge and presentations than the original unit which was proposed but proved unsatisfact ry. Secondly, the scale of the study was small and differences between teachers might vitiate the effects of the ability displayed on the learning potential task. Thirdly, a teacher was required to teach the regular class sections since these students were taken from study hall sections. Consequently, a teacher who was experienced in teaching science to students in a laboratory format was hired.

All the classes were taught by this teacher, Mrs. Jean Rosner. An assistant, Mrs. Karen Corbett Howell, was assigned to manage the materials distribution, and to work with the students requiring additional aid. This reduced the teacher-pupil ratio to optimum proportions (1 to 6). Classes ran forty-five minutes each, twice a week, for 13 weeks, January 9th to April 14th, 1967. In all, each class received no more than 24 class periods because of "snow days" when school was closed.

During the course of the intervention, an individual interview was conducted with each child in order to gain some sense of the child's perception of himself, his sense of commitment to, and responsibility for his efforts in school, his activities and interests outside of school, and the extent to which he felt he was able to influence the individuals in his environment and his own efforts. The relationship of some of these variables to his performance on electricity evaluation tasks will be presented in a subsequent section.

In addition to the posttest evaluation, an observation schedule was set up to try to capture the richness of the childrens' behaviors in class. It was hoped that a record of such behaviors would provide good information on how different EMRs and non-EMRs functioned in class and would help explain unusually wide ranges of gain from pre- to posttest among children of the same class status or learning potential. The students in each class were rated by two trained observers on their Motivation (evident interest in the work), Attention Span, Need for Acknowledgement, and Role in class. Ratings for the first three measures were scaled on a 3 (low) to 1 (high) forced choice scale. For overall Role in class, the students were classified as active-constructive, passive-constructive, undistinguished by either constructive or destructive behavior, active-disruptive, or passive-disruptive, (aiding or approving of active disrupter). This role classification has been dropped from the analysis because it did not lend itself to a linear scoring system. (Copies of these variable scales are presented in Appendix D). These scales were taken at every class for all children. In addition, teacher's ratings of children on the variables of Ability, Productivity and Application to work (or work accomplished), and Cooperation were collected biweekly.

IV. RESULTS

A. Comparisons of Electricity Test Scores.

Comparison of the pretest, posttest, and gain scores from the taught and nontaught special class samples served as the test of the hypotheses that the teaching intervention was effective, and that learning potential status would predict actual learning status following teaching. That is, the high scorers and gainers would demonstrate greater understanding of the simple concepts of electricity and obtain higher scores on the posttest than nongainers.

Comparison of the pretest, posttest, and gain scores for the taught special and regular class samples served as the test of the hypotheses that regular class students would learn more than the special class students, but that learning potential status would predict level of performance among these groups. Lastly, comparison of the pretest, posttest, and gain scores of the taught special and low achieving regular class students tested the null hypothesis that there would be no between group differences in performance levels. The within group differences would be accounted for by learning potential status. These comparisons will be discussed separately.

The comparisons between EMR and nonEMR taught groups of different learning potential status were performed by unequal N analysis of variance. This was necessary because EMR status and learning potential status are not independent of each other. High scorers are over-represented in the nonEMR group, whereas gainers and nongainers are over-represented in the EMR group. The Pearson r between EMR status and learning potential status is .24 ($p<.06$) if one assumes a linear equal interval continuum from nongainer to gainer to high scorer. Computation of the main effects of EMR status assume that there are the same proportions of nongainers, gainers, and high scorers in each of the two groups thus partialing out learning potential status. Computation of the main effect of learning potential status likewise assumes that EMRs and nonEMRs are represented in the same proportions within each learning potential status, thus partialing out the confounding effects of EMR status. The conclusions thus have to be modified somewhat by the fact that learning potential status and EMR status are not completely independent and have 6% of each others' variance in common.

Nine scores were derived from the evaluation instrument. The Boards presentation provided four scores - Simple Circuits, Complex Circuits I, Complex Circuits II, and the Boards Subtotal for this type of presentation. Four scores were derived from the administration of the Diagram section of the evaluation instrument. (Simple Circuits, Complex Circuits, Schematic Diagrams, Diagrams Subtotal). The ninth score was the total score for the test (the sum of the two subtotals).

1. Comparison of Taught and Nontaught Special Class Students.

a. Analysis of pretest scores.

Tables 2 and 3 present the nine mean pretest scores for the taught and nontaught special class samples, respectively. Both groups were subdivided by sex and learning potential status. Analyses of variance, (sex, 2 levels; learning potential status, 3 levels; taught - non-taught, 2 levels), were performed independently on the nine sets of pretest scores revealed no main effect for teaching. A non-significant multivariate F-ratio was obtained also suggesting the initial equality of the taught and non-taught samples (see Table 4).

b. Analysis of posttest scores.

The mean posttest scores for the subgroups in each sample, subdivided by sex and learning potential status, are presented in Tables 5 and 6 for taught and nontaught special class children, respectively. Table 7 summarizes the obtained F-ratios, for the nine posttest measures, and the multivariate F-ratios based on the nine dependent measures. As was predicted, the teaching main effect for five of the nine test scores was significant, as was the multivariate F-ratio, indicating that exposure to the curriculum unit did teach the simple concepts of electricity, These results support the first hypothesis regarding the efficacy of the teaching intervention.

Subjects of different learning potential status had significantly different mean scores on all but one of the nine test scores and there was a significant multivariate F-ratio for this variable. On all the measures, the performance of the high scorers was superior to that of the gainers, who in turn out-scored nongainers. The multivariate F-ratio associated with this linear continuum is 4473, $p<.001$. With the exception of interactions between learning potential status and teaching on Simple Circuits (Boards and Diagrams), no other effects in the analysis achieved significance. However when the degrees of freedom associated with learning potential were analyzed for a linear and quadratic component (1 df each), the linear trend for learning potential status accounted for all the significant effects, and also indicated significant interactions for two additional test scores (subtotal, Boards, and total test score). The results provide support for the second hypothesis that learning potential status would predict actual learning status, and the findings for the linear component indicate that learning occurred in the predicted order, i.e., high scorers were highest, then gainers and nongainers. Figure 2 and 3 present these results by learning potential status for the Boards and Diagrams subtotals, respectively, for the pre- and posttest scores.[1]

---

1. Because the learning potential groups differed in IQ levels, the

Table 2. Mean Pretest Scores on Electricity Test for Taught Special Class Students

	Simple Circuits	Complex Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Total
High Scorers-Male-(N=7)									
Means	2.43	10.14	4.00	16.57	3.86	10.21	5.86	19.86	36.43
S.D.	1.40	0.44	1.51	2.19	2.95	2.27	2.36	6.10	7.35
High Scorers-Female-(N=2)									
Means	2.00	11.50	3.00	17.00	1.00	11.00	6.00	18.00	34.50
S.D.	2.00	1.00	--	1.00	--	1.50	--	2.00	0.50
Gainers - Male (N=7)									
Means	3.57	9.50	3.29	16.29	3.57	8.07	6.14	17.71	34.00
S.D.	1.76	2.55	1.03	2.60	1.92	1.76	2.17	4.62	4.84
Gainers - Female (N=5)									
Means	2.60	8.20	3.00	13.80	1.20	6.44	5.00	12.60	26.60
S.D.	0.80	2.40	1.10	3.54	0.40	1.94	1.90	3.61	1.85
Nongainers - Male (N=4)									
Means	1.50	10.92	4.50	17.00	2.75	6.38	5.00	14.25	31.25
S.D.	1.12	1.52	2.06	2.35	0.83	1.75	0.71	2.49	3.49
Nongainers - Female (N=8)									
Means	2.25	8.10	2.50	12.75	2.00	7.01	5.88	14.88	27.63
S.D.	1.20	3.03	1.32	3.46	1.00	2.90	1.76	3.48	5.22

Table 3: Mean Pretest Scores on Electricity Test for Nontaught Special Class Students

	Simple Circuits	Complex Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Total
High Scorers-Male (N=4)									
Means	1.50	10.63	5.00	17.00	3.25	10.25	7.50	21.00	38.00
S.D.	0.50	2.33	1.41	2.35	1.30	2.05	2.60	5.24	7.21
High Scorers-Female (N=2)									
Means	2.50	10.00	4.50	17.00	1.50	12.00	6.50	20.00	37.00
S.D.	1.50	--	0.50	1.00	0.50	3.00	1.50	1.00	2.00
Gainers - Male (N=6)									
Means	2.00	10.75	4.33	17.00	3.33	8.67	6.00	17.67	35.00
S.D.	1.29	1.82	1.37	1.41	1.49	2.84	2.58	5.62	6.38
Gainers - Female (N=2)									
Means	1.50	9.50	3.00	14.00	3.00	7.25	5.50	16.00	30.00
S.D.	0.50	1.50	--	2.00	1.00	3.25	1.50	4.00	6.00
Nongainers - Male (N=5)									
Means	2.40	9.20	2.80	14.40	2.40	7.90	5.60	16.00	30.40
S.D.	1.02	2.14	0.40	2.87	1.02	1.96	2.42	4.24	4.32
Nongainers - Female (N=7)									
Means	3.86	8.43	3.43	15.71	1.71	8.10	4.86	14.57	30.43
S.D.	2.03	2.26	0.90	3.88	1.16	3.84	1.25	3.70	5.26

Table 4. Summary of Analysis of Variance (F-Ratios) of Pretest Scores for Taught and Non Taught Special Class Students.

	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total	Multivariate F-Ratio
Taught vs. Non Taught EMRs (T)	1	0.001	0.174	1.772	0.697	0.048	1.030	0.065	0.364	1.024	0.574
LP Status	2.	0.592	2.024	2.942	2.361	1.551	6.420[2]	1.015	5.117[1]	6.882[2]	1.790
Sex	1	0.392	3.016	2.867	2.863	9.448[2]	0.018	0.284	1.887	4.156[1]	1.771
T x LP	2	3.775[1]	0.683	0.949	0.017	0.363	0.078	0.709	0.080	0.040	0.888
T x Sex	1	1.460	0.104	1.514	1.889	1.176	0.060	0.256	0.101	0.977	0.476
LP x Sex	2	1.590	0.845	0.031	0.618	0.846	0.940	0.293	0.602	0.831	1.648
T x LP x Sex	2	0.084	0.709	2.173	1.450	0.396	0.058	0.311	0.371	0.051	0.825
Error Mean Square	47	2.408	5.635	1.846	9.661	2.879	8.237	4.950	23.918	34.999	11.049

[1] $p<.05$

[2] $p<.01$

Table 5. Mean Posttest Scores on Electricity Test for Taught Special Class Children

	Simple Circuits	Complex Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Total
High Scorers-Male (N=7)									
Means	7.71	13.43	6.14	27.29	7.71	13.43	10.29	31.57	58.71
S.D.	3.19	3.26	1.46	6.36	3.33	4.55	2.71	9.07	15.24
High Scorers-Female (N=2)									
Means	9.50	17.50	7.00	34.00	10.00	15.75	13.50	39.50	73.00
S.D.	2.50	0.50	1.00	3.00	2.00	0.75	0.50	2.50	5.00
Gainers - Male (N=7)									
Means	5.43	14.36	5.43	25.29	7.14	13.00	8.86	27.29	52.57
S.D.	2.77	3.40	1.50	5.77	3.00	4.43	2.95	9.94	14.85
Gainers - Female (N=5)									
Means	5.20	11.70	4.40	21.20	4.20	9.60	7.60	21.20	42.80
S.D.	1.72	3.01	1.20	4.96	1.60	1.59	2.80	4.62	7.47
Nongainers - Male (N=4)									
Means	3.25	10.63	3.00	17.00	3.00	12.38	6.75	22.00	39.25
S.D.	0.43	3.73	1.22	4.64	1.22	5.56	0.83	6.00	9.73
Nongainers - Female (N=8)									
Means	3.50	9.38	3.38	16.25	4.25	8.13	5.13	17.50	33.63
S.D.	1.00	3.34	1.49	4.44	1.09	3.26	2.57	5.17	8.99

Table 6. Mean Posttest Scores on Electricity Test for Nontaught Special Class Children

	Simple Circuits	Complex Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Total
High Scorers - Male (N=4)									
Means	2.25	10.75	4.25	17.25	3.00	9.50	8.25	21.00	38.25
S.D.	1.09	3.09	1.09	4.76	1.58	1.54	1.30	2.12	6.10
High Scorers - Female (N=2)									
Means	5.00	11.25	4.00	20.50	1.50	13.25	8.50	23.00	43.50
S.D.	3.00	1.25	1.00	5.50	1.50	2.25	0.50	3.00	8.50
Gainers - Male (N=6)									
Means	3.33	11.25	4.50	20.67	3.00	9.50	6.67	17.67	40.17
S.D.	1.11	3.41	1.26	2.75	2.08	2.80	3.25	8.79	9.30
Gainers - Female (N=2)									
Means	3.00	10.75	3.00	16.50	2.00	10.25	8.00	20.00	37.00
S.D.	1.00	0.75	--	1.50	1.00	1.25	2.00	2.00	4.00
Nongainer - Male (N=5)									
Means	2.00	9.30	2.60	13.80	2.80	9.20	5.60	17.60	31.60
S.D.	0.63	1.72	0.80	2.79	1.72	1.75	2.73	4.50	6.28
Nongainer - Female (N=7)									
Means	2.43	10.79	2.86	16.00	2.29	8.23	6.29	17.43	33.71
S.D.	1.59	3.69	1.12	4.00	1.03	3.20	1.03	1.05	4.23

Table 7. Summary of Analyses of Variance (F-ratios) of Posttest Scores for Taught and Nontaught EMRs

	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total	Multivariate F-Ratio
Taught Groups (T)	1	18.093[2]	2.370	8.199[2]	11.070[2]	26.534[2]	3.609	3.201	9.547[2]	11.261[2]	4.062[2]
Learning Potential Status (LP)	2	11.322[2]	4.289[1]	16.015[2]	15.404[2]	4.047[1]	2.632	11.165[2]	9.972[2]	14.395[2]	3.141[2]
Sex	1	0.939	1.006	5.418[1]	3.328	3.950	3.586	1.828	3.107	4.291[1]	1.454
T x LP	2	3.432[1]	1.749	1.704	2.988	4.435[1]	0.462	1.645	2.528	2.676	1.066
T x Sex	1	0.050	0.305	0.335	0.029	0.669	2.157	0.613	0.657	0.310	1.077
LP x Sex	2	0.962	1.118	1.548	2.454	1.637	2.133	0.673	1.158	1.856	1.182
T x LP x Sex	2	0.055	0.847	0.119	0.375	1.457	0.112	1.142	0.882	0.610	0.708
Error Mean Square	47	4.532	12.097	1.957	26.661	5.134	14.076	7.039	50.228	12120.075	11.098

[1] $p < .05$

[2] $p < .01$

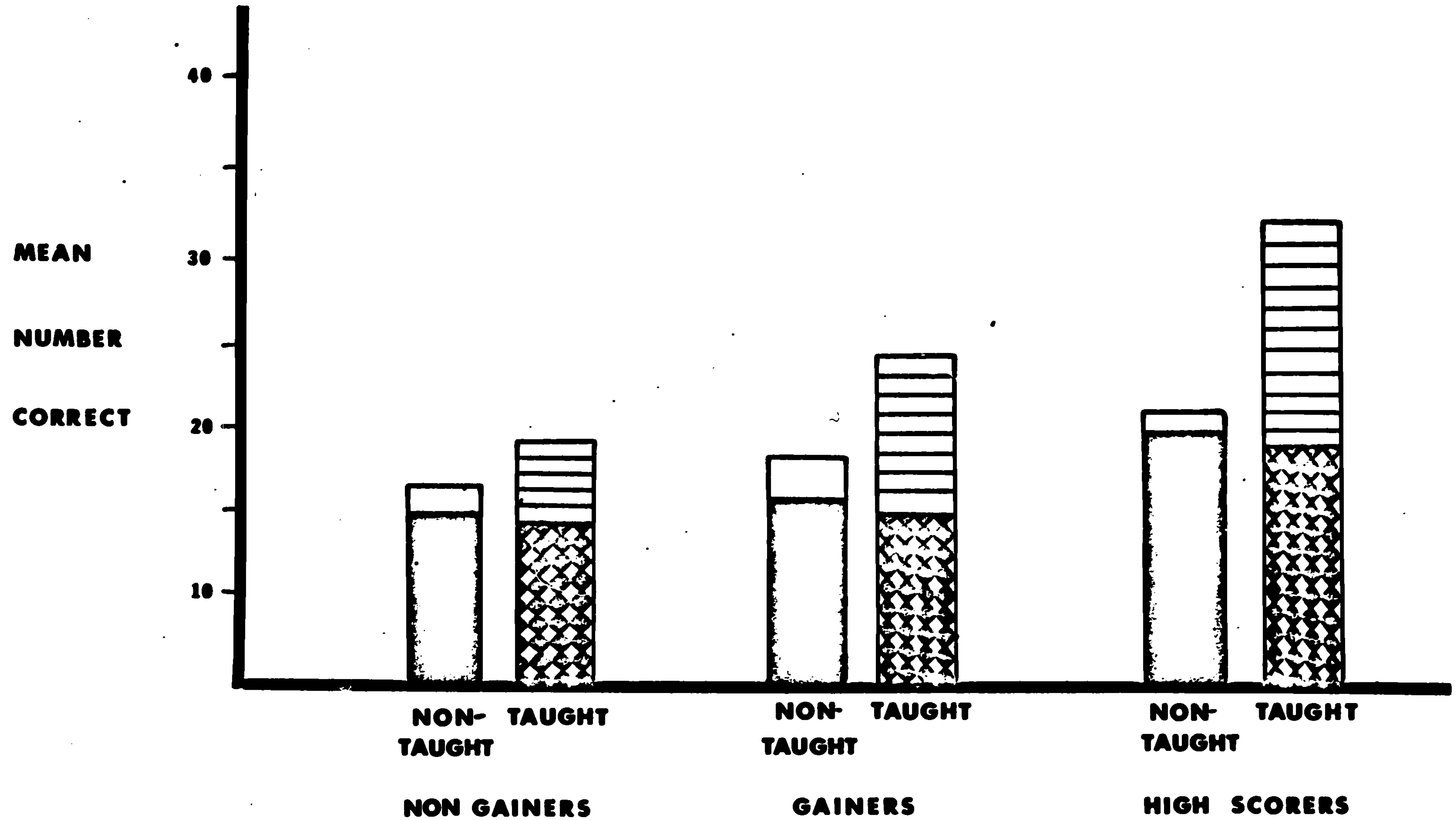


Figure 2. Comparison of Boards Subtotal Scores for Pre- and Posttest for Taught and Nontaught Special Class Students

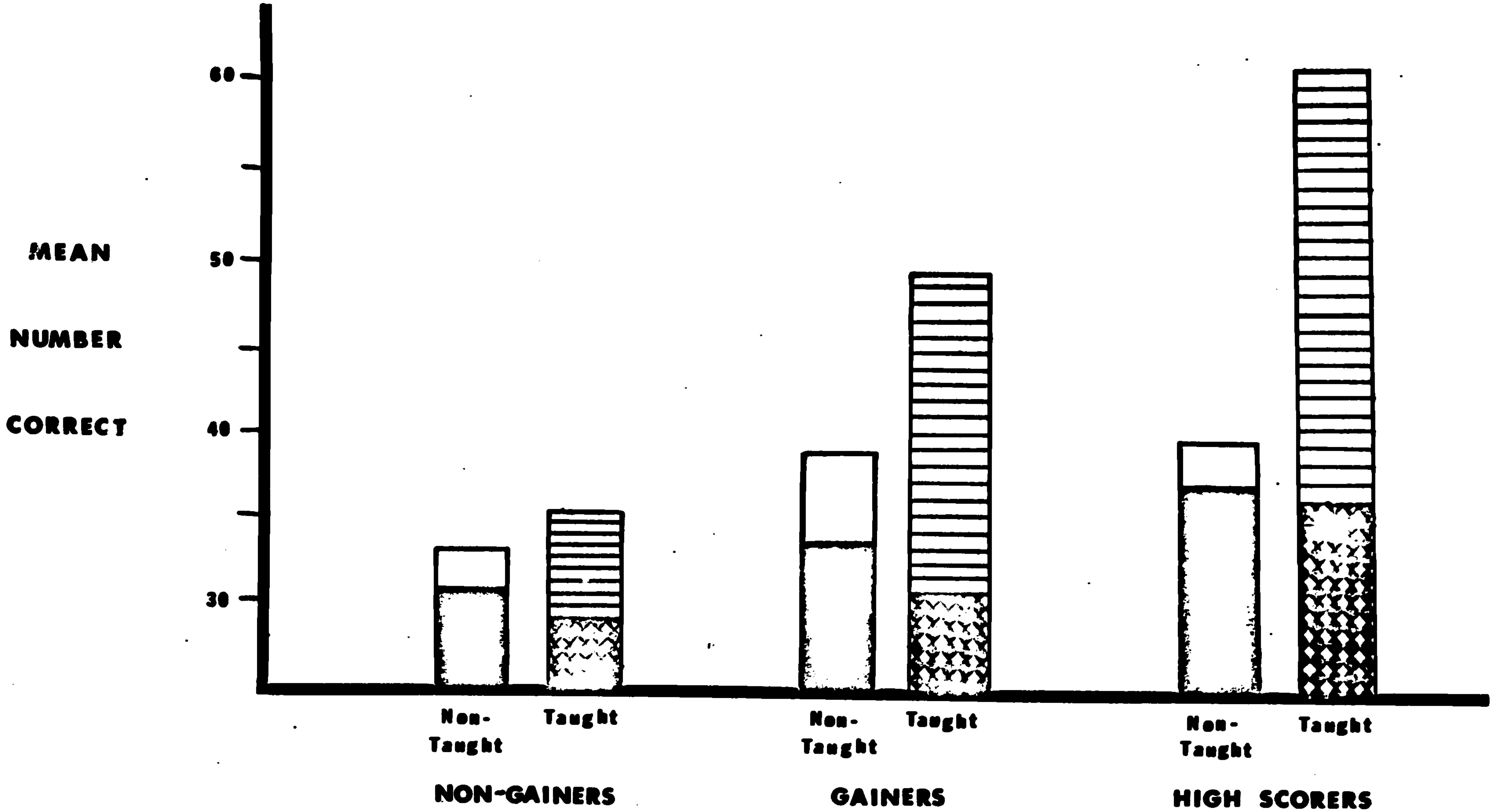


Figure 3. Comparison of Diagrams Subtotal Scores for Pre- and Posttest for Taught and Nontaught Special Class Students

c. Analysis of Improvement Scores[2]

The comparability of the EMR taught and non-taught groups having been established, a comparison of their mean improvement scores (posttest minus pretest) should also indicate the effects of the teaching intervention and learning potential status. Analyses of variance of gain scores were carried out separately for each of the nine test measures. Tables 8 and 9 present the mean gain scores for the taught and nontaught samples, respectively. Table 10 presents a summary of the obtained F-ratios. The most striking result is the reliable effect of the treatment factor for seven of the nine test variables. The teaching intervention failed to affect gain scores only on Complex Circuits 1 (Boards) and Schematic Diagrams. As with the posttest findings, these results indicate that the gain shown by the special class sample following teaching represented a real gain in their understanding of the electricity concepts taught, an increase not attributable to practice on the evaluation instrument. There were no significant sex differences.

While only three scores were significant for the main effect of learning potential status, the multivariate F-ratio was also significant, indicating as with the posttest scores, that subdividing the students by learning potential status accounted for a real proportion of the changes in score levels. The multivariate F-ratio for the linear component associated with the ranking of the groups (high scorer, gainer, and nongainer) was 4.166,($p<.001$), and six of the nine scores were significant in the univariate analyses for this component. No quadratic component effects were significant. This indicates a strong determining effect for the learning potential variable, according to hypothesized ability. To some extent, these differences in gain occurred regardless of whether they were taught, although as Figure 2 and 3 indicate, the taught high scorers and gainers did gain more markedly than their nontaught peers. The multivariate F-ratio for the linear component of the learning potential x taught - nontaught interaction was 1.951, ($p<.07$).Teaching made little difference for the nongainers since neither the taught or the nontaught group improved their scores.

---

1. scores were re-analyzed with an IQ covariate in order to assure that these IQ differences did not significantly determine the electricity test scores. The results were virtually unchanged. The taught special class students learned from the classes. The variance attributable to the learning potential variable was clearly due to the linear component (seven of the nine F-ratios and the multivariate F-ratio were significant, $p< .002$). There were no significant F-ratios for the quadratic component. Similarly the taught-nontaught x learning potential interaction, linear component, accounted for all the significant effects.

2. The improvement scores were increased by 50 to eliminate negative scores.

Table 8. Mean Gain Scores on Electricity Test for Taught Special Class Children

	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total
High Scorers-Male (N=7)									
Means	55.29	53.29	52.14	60.86	53.86	53.21	54.43	61.57	72.00
S.D.	2.81	2.91	1.36	5.11	2.23	4.35	3.89	8.55	13.76
High Scorers-Female (N=2)									
Means	57.50	56.00	54.00	68.00	59.00	54.75	57.50	71.00	88.50
S.D.	4.50	1.50	1.00	4.00	2.00	0.75	0.50	1.00	4.50
Gainers-Male (N=7)									
Means	51.86	54.86	52.14	59.00	51.86	54.84	54.14	57.14	64.43
S.D.	3.68	5.12	1.12	6.39	3.44	4.76	8.24	10.60	13.68
Gainers-Female (N=5)									
Means	52.60	53.50	51.40	57.80	53.00	53.16	52.60	58.80	66.00
S.D.	2.06	1.58	1.85	4.92	1.26	3.33	2.73	5.74	9.25
Nongainers-Male (N=4)									
Means	51.75	49.70	48.50	50.00	50.25	56.00	51.75	58.25	57.50
S.D.	0.83	3.22	2.96	4.74	1.92	6.84	0.83	8.07	9.94
Nongainers-Female (N=8)									
Means	51.25	51.27	50.88	53.38	52.25	51.11	49.25	52.63	55.63
S.D.	1.92	3.29	2.71	5.98	1.85	2.95	2.33	5.00	7.97

Table 9. Mean Gain Scores on Electricity Test for Nontaught Special Class Children

	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total
High Scorers-Male (N=4)									
Means	50.75	50.13	49.25	50.50	49.75	49.25	50.75	49.75	50.25
S.D.	0.83	2.56	1.09	3.64	1.92	0.56	2.86	4.49	1.48
High Scorers-Female (N=2)									
Means	52.50	51.25	49.50	53.50	50.00	51.25	52.00	53.50	56.50
S.D.	1.50	1.25	1.50	4.50	2.00	0.75	1.00	2.50	6.50
Gainers - Male (N=6)									
Means	51.33	52.08	50.17	53.50	49.67	50.83	50.67	51.00	55.00
S.D.	1.60	2.03	1.86	3.45	2.69	2.36	2.69	5.03	7.39
Gainers - Female (N=2)									
Means	51.50	51.25	50.00	52.50	49.00	53.00	52.50	54.50	57.50
S D.	0.50	0.75	---	0.50	---	2.00	0.50	1.50	1.50
Nongainers - Male (N=5)									
Means	49.60	50.10	49.80	49.60	50.40	51.38	58.20	51.80	51.40
S.D.	1.02	3.67	1.17	5.39	1.50	2.29	16.45	5.95	9.13
Nongainers-Female (N=7)									
Means	48.57	52.36	49.43	50.43	50.57	50.90	51.29	52.86	53.14
S.D.	1.05	4.03	1.76	3.37	1.92	3.24	1.67	3.56	3.14

Table 10. Summary of Analysis of Variance (F-Ratios) of Gain Scores for Taught and Non Taught Special Class Students.

	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total	Multivariate F-Ratio
Taught vs. Non Taught EMRs	1	16.609[2]	2.621	10.234[2]	17.249[2]	18.521[2]	5.997	0.008	11.073[2]	17.973[2]	4.186[2]
LP Status	2	9.713[2]	1.854	2.515	7.386[2]	2.644	0.330	0.202	1.837	5.419[2]	2.046[1]
Sex	1	0.041	0.612	0.458	0.802	2.836	0.662	0.975	0.255	0.799	0.920
T x LP	2	2.379	1.752	2.157	3.145	3.535	0.263	1.969	2.536	4.398	1.499
T x Sex	1	0.241	0.076	1.046	0.234	2.998	1.759	0.241	0.175	0.009	1.230
LP x Sex	2	1.217	1.032	1.090	1.322	1.039	1.366	1.118	1.585	1.318	0.855
T x LP x Sex	2	0.000	0.102	0.781	0.144	0.439	0.245	0.393	0.699	0.449	0.405
Error Mean Square	47	5.963	13.291	4.115	29.531	5.936	15.907	45.007	51.945	104.658	10.971

[1] $p<.05$

[2] $p<.01$

The main conclusions to be drawn from the foregoing comparisons of the taught and non-taught EMR samples are:

1. That EMR children exposed to a unit in electricity taught with manipulative techniques do gain in their knowledge of the subject, (as measured by the present testing instrument);

2. That differential rates of improvement of understanding following the learning experience are evident within the "homogeneous IQ" EMR group;

3. That a substantial portion of these differences in gain may be predicted from their learning potential status classifications, high scorers gaining more than gainers who, in turn, improved more than nongainers.

2. Comparison of Taught Special and Regular Class Students.

The third and fourth hypotheses were concerned with the relative ability of special and regular class samples of students to benefit from the Batteries and Bulbs course and the role of learning potential status in influencing the scores systematically.

a. Analysis of Pretest Scores.

The effects of school status (special vs regular), sex, and learning potential status (3) were determined by analyses of variance of the nine separate dependent scores. A multivariate F-ratio was obtained for each effect based on all nine scores.

Tables 2 and 11 present the mean pretest scores for the special and regular class students, respectively. Table 12 summarizes the F-ratios obtained in the analyses of variance of the pretest scores. It is evident that the differences in class status influenced pretest score levels, as the scores on six of the nine dependent measures, including both subtotals and total score were significant, as was the multivariate F-ratio for this main effect. As predicted, regular class students scored higher than the special class students on the pretest. Learning potential status minimally influenced the pretest scores, (significant main effect only for Complex Circuits (Diagrams)). There were some indications of sex differences in performance,---main effect, subtotal (Boards) and total score; sex x class interactions for Complex Circuits (Boards), and the sex x learning potential status x school status interaction.

b. Analysis of posttest scores.

Since, in general, the regular class students demonstrated superior knowledge of electricity on the pretest as was hypothesized, it may be argued that one can not compare the posttest and gain scores of the two samples in order to determine whether they benefitted differentially from the intervention. The possibility exists that a ceiling effect may have depressed the posttest scores of the initially superior regular class group.

Table 11. Mean Pretest Scores on Electricity Test for Taught Regular Class Students

	Simple Circuits	Complex Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Total
High Scorers - Male (N=10)									
Means	3.10	11.25	4.00	18.30	3.50	10.15	7.90	21.60	39.90
S.D.	2.17	2.88	1.18	3.69	2.46	1.57	1.45	3.26	6.07
High Scorers - Female (N=7)									
Means	2.43	10.71	3.86	16.86	3.29	10.50	7.14	21.00	38.00
S.D.	1.40	2.27	0.99	3.31	1.48	1.41	1.81	2.73	4.50
Gainers - Male (N=2)									
Means	4.50	9.00	2.00	15.50	4.50	7.00	4.00	15.50	31.00
S.D.	0.50	2.00	1.00	3.50	1.50	1.00	1.00	3.50	7.00
Gainers - Female (N=6)									
Means	3.17	9.88	4.00	17.00	3.67	9.42	6.67	19.83	36.83
S.D.	1.77	1.35	0.82	1.53	1.49	1.97	1.49	3.58	3.98
Nongainers-Male (N=5)									
Means	3.20	12.70	2.60	18.40	3.00	10.30	8.40	21.40	40.40
S.D.	1.60	2.14	1.20	1.50	1.55	1.69	2.42	4.45	5.82
Nongainers - Female (N=3)									
Means	3.00	11.17	3.33	17.67	4.00	9.00	5.67	18.67	36.33
S.D.	[illegible].16	2.39	1.25	1.25	0.82	5.34	2.87	6.80	7.41

2

Table 12. Summary of analysis of Variance (F-Ratios) of Pretest Scores for Taught Special and Regular Class Students

	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total	Multivariate F-Ratio
pecial & Regular Class	1	1.561	6.490[1]	0.528	8.904[2]	3.342	9.156[2]	8.212[2]	14.130[2]	20.864[2]	2.500[1]
P Status	2	1.494	1.458	2.091	0.833	0.319	4.957[1]	0.960	3.068	3.132	1.139
ex	1	1.290	3.064	0.294	5.038[1]	3.278	0.097	0.697	1.596	4.974[1]	1.126
lass x LP	2	0.106	0.969	0.417	0.624	0.529	2.087	0.628	0.650	1.160	0.797
lass x Sex	1	0.263	0.363	5.015[1]	1.250	3.237	0.225	0.180	0.800	1.614	1.081
P x Sex	2	0.792	0.966	1.947	0.399	1.042	0.719	0.149	0.065	0.046	0.716
lass x LP x Sex	2	0.050	0.794	0.581	1.180	0.098	1.666	3.431[1]	2.131	1.795	1.462
rror Mean Square	54	3.146	6.265	1.836	9.829	3.813	5.997	4.251	20.421	34.512	10.713

$p<.05$

$p<.01$

However, it should be noted that the mean pretest scores of this group were far below the 50% correct level on the test, so that there can be no objection that there was "nowhere for them to go". This does not nullify the possibility that there may be different degrees of difficulty in improving a score when the starting levels differ markedly. Therefore, the qualifying fact that the groups differed in pretest level must be considered in interpreting differences in posttest and gain scores.

Tables 5 and 13 present the mean posttest scores for the special and regular class students respectively. Table 14 summarizes the F-ratios obtained. As hypothesized, the regular class students did tend to do better on the subtests, and on the Complex Diagram subtests in particular. The differences due to class placement were less pervasive after teaching the unit. There were fewer significant F-ratios, those that were significant had lower F-ratios, and the multivariate F-ratio was no longer significant. There were instances for some scores in which the special class students' means were slightly higher, as comparison of the tables of mean scores indicates.

The more interesting differences occurred in examining the only other major effect influencing the posttest scores systematically, namely, learning potential status. This main effect was highly significant for eight of the nine scores, and also yielded a significant multivariate F-ratio ($F=2.391, p<.01$), in sharp contrast to the lack of effect evident on the pretest analyses. As Figures 4 and 5 indicate, high scorers tended to score highest, then gainers, with nongainers scoring lowest irrespective of class placement. This was particularly true for the special class student, since the regular class students tended to show less dramatic differences by learning potential. The multivariate F-ratio for this linear continuum was $3.236, p<.004$.

Analysis of the linear and quadratic components of the learning potential variable indicated a significant effect for the triple interaction of taught groups x sex, x learning potential, quadratic component for four scores. Plotting the means, this triple interaction indicated that the female regular class gainers scored equally high as the female high scorers. The male regular class gainers scored lower than their nongainer and high scorer peers. While learning potential status was generally predictive of score levels attained, the differences among the students in three regular class groups were not as dramatic as they were among the special class students, nor was the predicted linearity of dimension invariable maintained.[3]

---

3. The posttest scores were re-analyzed in the same design covarying for the differences in IQ, which adjusted the scores for the differences between the two taught samples, and for the learning potential groups within each sample. Adjusting these scores for IQ level produced some interesting changes in the pattern of F-ratios. The small

Table 13. Mean Posttest Scores on Electricity Test for Taught Regular Class Students

	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total
High Scorers-Male (N=10)									
Means	7.10	15.20	5.90	28.30	8.00	13.10	10.70	31.70	60.00
S.D.	3.14	2.82	3.94	5.40	3.32	3.73	2.72	7.46	10.39
High Scorers-Female (N=7)									
Means	6.14	14.86	5.71	26.86	6.00	14.43	11.57	32.29	58.57
S.D.	3.09	3.70	1.67	6.42	3.89	3.91	2.13	7.52	13.16
Gainers-Male (N=2)									
Means	3.50	12.50	6.00	22.00	4.00	14.00	7.50	26.00	47.50
S.D.	1.50	2.00	1.00	2.00	1.00	1.50	0.50	2.00	3.50
Gainers-Female (N=6)									
Means	3.83	18.08	5.00	26.83	6.17	14.83	11.00	32.00	59.00
S.D.	2.73	4.31	1.53	6.39	3.13	4.30	4.32	9.59	15.86
Nongainers-Male N=5)									
Means	6.00	14.10	4.20	24.20	5.80	15.00	11.00	32.00	56.00
S.D.	3.16	2.99	1.17	6.46	3.19	4.18	4.47	11.01	17.45
Nongainers-Female (N=3)									
Means	4.00	11.17	3.00	18.00	3.33	9.17	5.67	18.33	36.33
S.D.	0.82	3.47	0.82	4.90	0.47	5.10	1.25	6.13	11.03

Table 14. Summary of analysis of Variance (F-Ratios) of Posttest Scores for Taught Special and Regular Class Students.

	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total	Multivariate F-Ratio
Special vs. Reg.Class	1	0.122	8.776[2]	0.701	5.001[1]	0.365	4.168[1]	8.444[2]	6.449[1]	6.402[1]	1.8757
LP Status	2	7.940[2]	5.743[2]	7.267[2]	11.731[2]	6.449[2]	1.817	7.243[2]	6.393[2]	9.394[2]	2.394[2]
Sex	1	3.099	0.376	1.748	2.873	3.355	2.714	1.728	3.264	3.566	.778
Class xLP	2	1.947	0.904	0.253	1.244	0.403	1.008	1.306	1.518	1.509	0.918
Class x Sex	1	0.839	0.203	0.195	0.078	0.427	0.670	0.026	0.015	0.004	0.347
LP x Sex	2	0.137	0.697	0.492	0.618	0.138	2.495	2.983	2.110	1.487	1.340
Class x LP x Sex	2	0.414	3.435[1]	0.142	2.331	3.057	0.552	2.162	1.938	2.370	1.0154
Error Mean Square	54	7.931	13.370	4.866	37.329	9.350	19.106	9.826	73.605	189.546	10.65

[1] $p<.05$

[2] $p<.01$

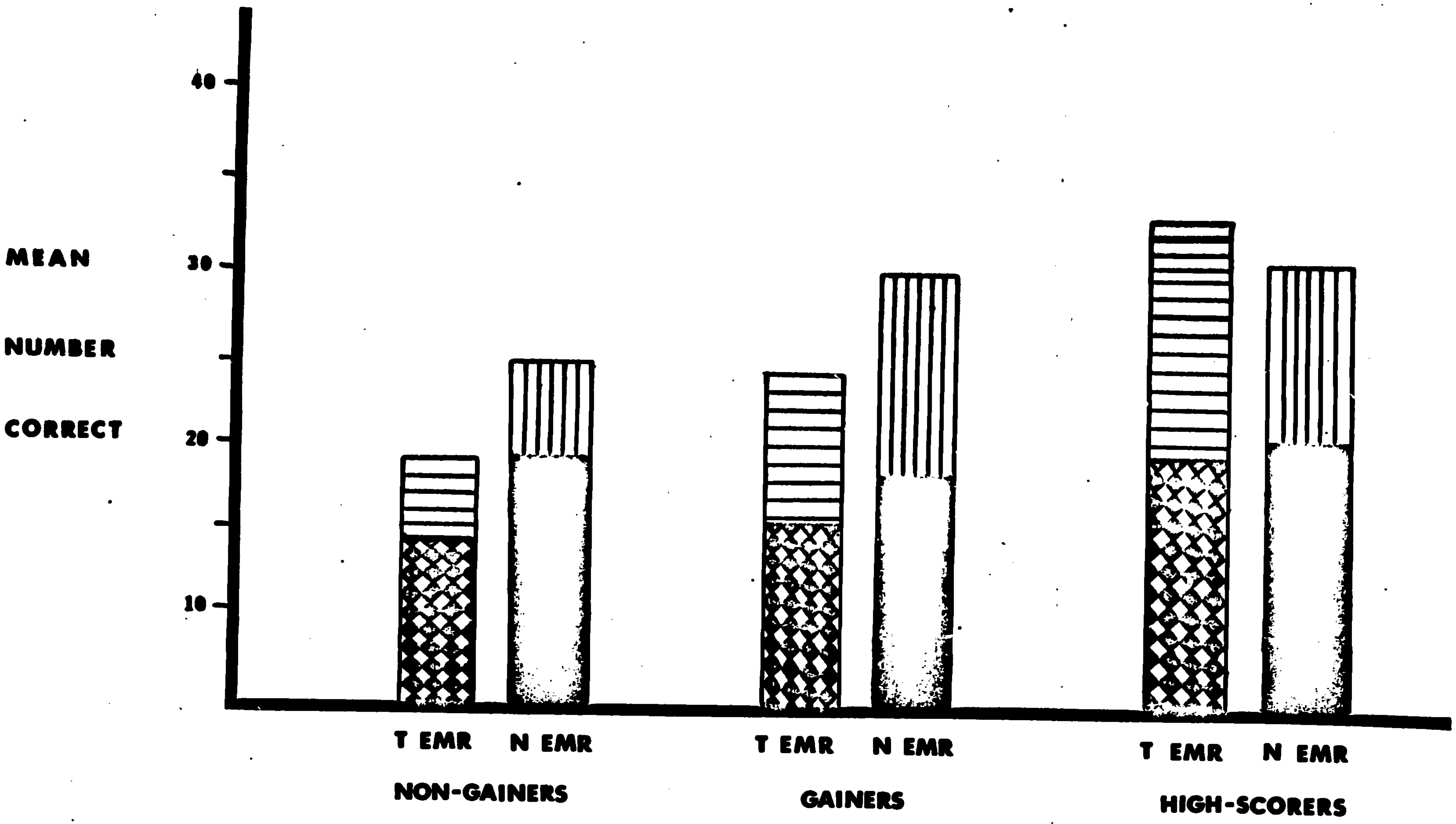


Figure 4. Comparison of Boards Subtotal Scores for Pre- and Posttest for Taught Special and Regular Class Students

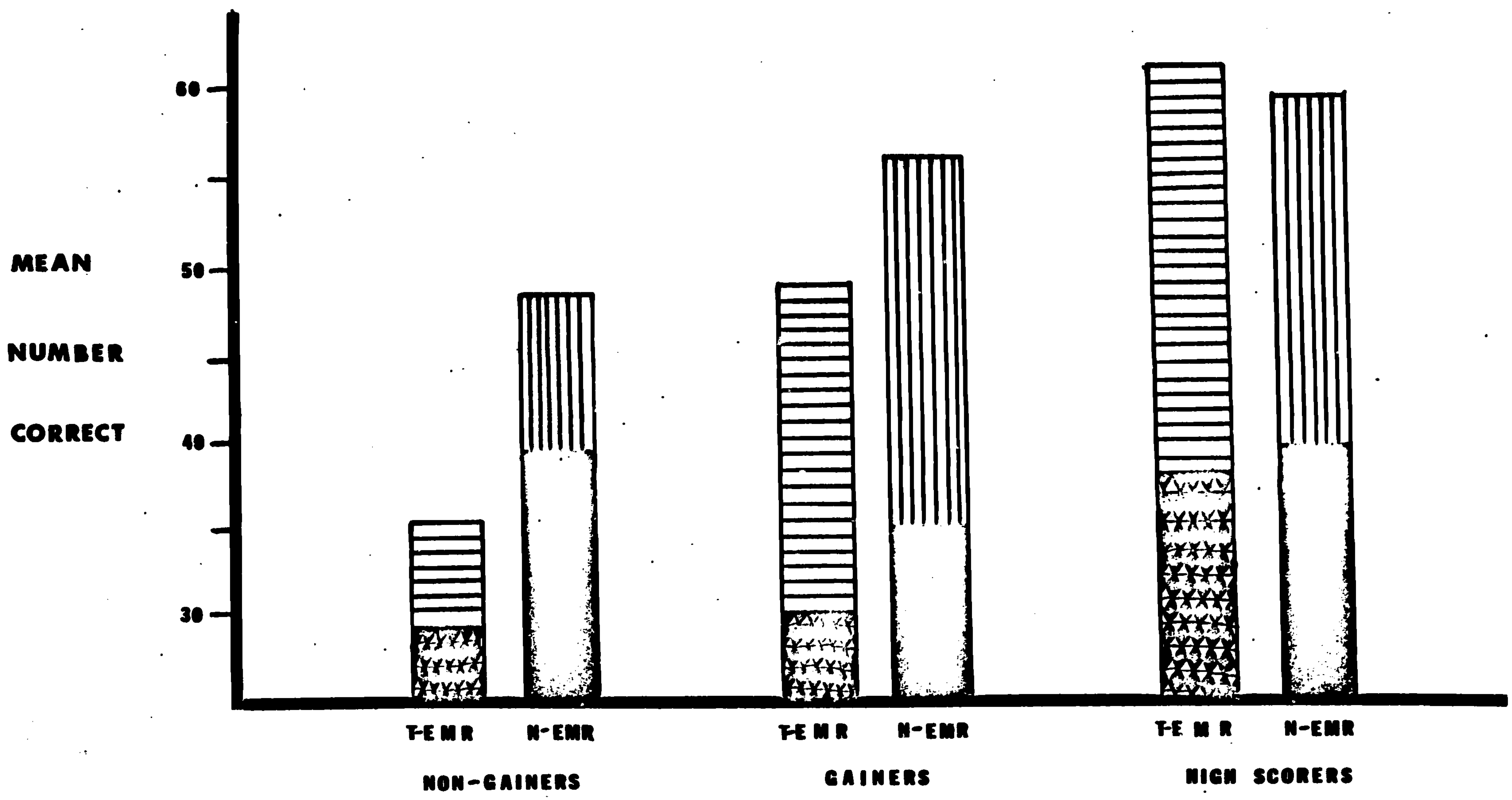


Figure 5. Comparison of Diagrams Subtotal Scores for Pre- and Posttest for Taught Special and Regular Class Students

There were no effects attributable to sex differences, thus, although the results tend to support the hypothesis, the superiority of the regular class students which was evident on the pretest was not marked, in absolute score level, on the posttest.

c. Analysis of improvement scores.

Tables 8 and 15 present the mean gain scores for the taught special and regular class students. Table 16 summarizes the F-ratios obtained in the separate analyses of variance for each of the nine test-derived scores, and presents the multivariate F-ratio for each effect.

When the improvement scores are compared, the trend evident in the posttest comparisons is clearer. As Table 16 indicates, only the main effect of learning potential status differentiates consistently among the students. Only one interaction, (class x sex, Simple Circuits, Diagrams), was significant for the class placement variable. By contrast, learning potential status was significant for five of the scores, and yielded a significant multivariate F-ratio. The linear component of the multivariate F- ratio was 3.569 ($p<.002$) and the scores for eight of the nine measures were significant for the univariate analyses (all at $<.002$). The effect of learning potential was highly determinating of degree of improvement following teaching, particularly when the groups were ranked according to hypothesized ability, as in the analyses of the linear component.

When relative improvement from pre- to posttest was analyzed there was no effect of class placement on any of the gain scores. Therefore the null hypothesis that special class children gained from the unit to the same extent as their regular class counterparts may not be discarded.

In summary, analysis of the gain scores expecially, suggests the rather striking conclusion that grouping students by Kohs learning potential status, rather than by IQ or school placement status, provided a better prediction of their ability to profit

---

3. (continued) number of significant differences between the two taught groups disappeared, there being no significant F-ratios for the main effect of groups. However, the pronounced sex differences in pretest scores remained following teaching, (six of nine significant F-ratios). The quadratic component for the triple interaction mentioned above also appears, and a similar explanation accounts for the patterns of mean scores. The consistently significant learning potential effect was largely accounted for by the linear component (five of nine scores and a significant multivariate F-ratio).

Table 15. Mean Gain Scores on Electricity Test for Taught Regular Class Students

	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total
High Scorers-Male (N=10)									
Means	54.00	53.95	51.90	58.90	54.50	52.95	52.80	60.30	70.10
S.D.	3.22	3.98	4.09	7.34	3.56	3.45	2.79	6.66	11.05
High Scorers - Female (N=7)									
Means	53.71	54.14	51.86	59.57	52.71	53.93	54.43	60.86	70.71
S.D.	3.69	3.92	1.81	6.65	4.23	2.92	2.72	5.51	9.97
Gainers - Male (N=2)									
Means	49.00	53.50	54.00	57.00	49.50	57.00	53.50	60.00	66.50
S.D.	1.00	4.00	2.00	5.00	2.50	2.50	0.50	6.00	10.50
Gainers - Female (N=6)									
Means	50.67	58.20	51.00	60.00	52.50	55.42	52.67	62.33	72.00
S.D.	2.87	4.64	1.41	6.53	3.20	5.50	2.87	12.04	18.40
Nongainers-Male (N=5)									
Means	52.80	51.40	51.60	55.80	52.80	54.70	52.60	60.00	65.80
S.D.	3.87	2.40	1.20	5.15	1.83	3.37	4.18	7.59	12.30
Nongainers-Female (N=3)									
Means	51.00	50.00	49.67	50.67	49.33	50.17	50.00	49.33	50.00
S.D	2.94	2.16	1.70	5.79	0.94	3.17	2.83	4.78	8.64

Table 16. Summary of Analysis of Variance (F-Ratios) of **Gain Scores for Taught Special and Regular Class Students.**

	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total	Multivariate F-Ratio
Special and Regular Class	1	0.137	1.249	0.111	0.137	0.006	0.128	0.031	0.570	0.948	1.062
LP Status	2	7.191[2]	6.810[2]	2.631	7.460[2]	5.093[2]	1.088	3.771[1]	3.138	6.639	2.602[2]
Sex	1	0.036	0.658	0.097	0.150	0.239	1.428	0.387	0.083	0.004	0.934
Class x LP	2	1.022	0.577	0.385	0.923	0.255	0.404	0.645	0.968	1.086	0.746
Class x Sex	1	0.253	0.001	2.407	0.578	4.861[1]	0.297	0.196	0.325	0.790	1.130
LP. x Sex	2	0.501	0.023	1.697	0.492	0.435	2.062	1.545	2.611	1.688	0.827
Class x LP x Sex	2	0.286	1.676	0.278	1.085	2.372	0.013	0.071	0.337	0.725	0.592
Error Mean Square	54	11.034	15.892	6.675	44.178	9.655	19.695	17.683	71.870	169.569	10.665

[1] $p < .05$

[2] $p < .01$

from the Batteries and Bulbs unit, when gain from the intervention was measured by the present measuring test instrument. However, as the following section indicates, two-thirds of the regular class sample consisted of low achieving students. The diminishing degree of difference between pre- and posttest scores in absolute score level, and the equivalent amount of gain may be due largely to this factor of sample composition. Hence, one can not conclude that special and adequately achieving regular class students' performances were equivalent.

3. Comparison of classroom learning by special and low achieving regular class students.

If, as the learning potential argument implies, special class high scorers and gainers appear to be more clearly educationally than mentally retarded, their performances should be similar to that of the low achieving regular class students who are educationally retarded since their IQs are in the dull and average ranges.

A test of this hypothesis was possible within the framework of this study. Within the regular class controls who participated in the classes, there was a considerable number of students who had achieved poorly in their major subjects, and who also tended to have IQs in the dull to average ranges. A sample of 22 regular class students met the criterion of a grade point average of less than 2.0 (less than a C average) based on the first four of five marking periods during the school year in which the electricity unit was taught. The characteristics of this sample, subdivided by learning potential status, is presented in Table 17, which also compares them to the taught special class samples. Differences in IQ level and CA were highly significant between the taught groups. The CA difference was due to the selection of 7th grade students for the regular class sample. There were also significant IQ differences among the learning potential groups within each sample. Nongainers had lower IQs than high scorers and nongainers. There was no significant difference in IQ between these latter groups within each sample. Both samples were drawn from similarly poor social backgrounds. The significant F-ratio for teaching groups is not a conceptually meaningful difference since all the subgroups fall in the lowest occupational categories. The nine dependent measures drawn from the scores on the pretest, posttest and gain scores were analyzed as described previously. The means for the pre-posttest and gain scores for the low achieving students are presented in Table 18. Tables 2, 5, and 8, respectively, present these scores for the special class students.

Reference to Table 19, which summarizes the analyses of variance of the pretest scores for the taught EMR and low achieving regular class students indicates that there were differences in knowledge of electricity prior to the start of the classes. These

Table 17.

Means, SDs & F-Ratios for Taught EMR and Low Achieving CA Control Samples for IQ, CA and Occupational Rating of Principal Wage Earner

	N	IQ $\bar{X}$	IQ SD	CA $\bar{X}$	CA SD	Mean Occupational Rating (Social Class Measures) $\bar{X}$	SD
EMR							
High Scorers	9	71.44	11.11	177.00	5.12	2.111	0.69
Gainers	12	73.00	5.29	170.83	10.44	2.167	0.76
Nongainers	12	68.17	5.31	165.17	9.91	2.083	1.14
Low Achieving CA Controls							
High Scorers	12	89.75	10.36	162.08	10.60	2.083	1.08
Gainers	4	90.25	2.50	163.75	8.77	.750	1.50
Nongainers	6	84.50	9.73	165.50	13.26	1.167	1.17

	df	IQ	CA	Social Class
Learning Potential Status (LP)	2	4.79^{1}	<1	<1
Teaching Group (G)	1	52.43^{2}	9.34^{2}	5.82^{1}
LP x G	2	<1	1.44	2.07
Residual Mean Square	49	66.906	100.924	.938

1 $p < .05$

2 $p < .01$

Table 18. Mean Scores on Electricity Test for Low Achieving Regular Class Students

	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total
High Scorers-Male (N=9)									
Pre	3.222	11.500	3.889	18.556	3.667	10.056	7.667	21.444	40.033
Post	6.889	14.667	6.000	27.667	7.667	12.667	10.556	30.778	58.944
Gain	53.667	53.167	52.111	57.889	54.000	52.611	52.889	59.556	68.444
High Scorers-Female (N=3)									
Pre	3.000	12.500	4.000	19.333	3.333	11.667	8.000	23.000	42.700
Post	8.667	17.333	6.333	32.333	9.667	17.167	12.667	40.000	72.200
Gain	55.667	54.833	52.333	62.667	56.333	55.500	54.667	66.333	79.333
Gainers-Male (N=1)									
Pre	5.000	11.000	3.000	19.000	6.000	8.000	5.000	19.000	38.000
Post	5.000	10.500	5.000	20.000	3.000	12.500	8.000	24.000	44.500
Gain	50.000	49.500	52.000	52.000	47.000	54.500	53.000	54.000	56.000
Gainers-Female (N=3)									
Pre	2.000	9.767	4.333	16.000	3.000	9.167	7.667	20.000	36.000
Post	1.667	15.667	6.000	23.333	3.333	14.000	9.333	26.667	50.433
Gain	49.667	55.900	51.667	57.333	50.333	54.833	51.667	57.000	64.000
Non-Gainers-Male (N=3)									
Pre	3.333	13.167	1.667	18.000	2.333	9.667	9.000	20.667	39.333
Post	5.000	12.667	3.667	21.333	4.000	13.667	9.000	27.000	48.500
Gain	51.667	49.500	52.000	53.000	51.667	54.000	50.000	55.667	59.000
Non-Gainers-Female (N=3)									
Pre	3.000	11.167	3.333	17.667	4.000	9.000	5.667	18.667	36.333
Post	4.000	11.167	3.000	18.000	3.333	9.167	5.667	18.333	36.800
Gain	51.000	50.000	49.667	50.667	49.333	50.167	50.000	49.333	50.000

Table 19. Summary of Analyses of Variance (F-ratios) of Pretest Scores for Taught EMRs and Low Achieving Regular Class Students

	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total	Multivariate F-Ratio
EMR vs. Low Achieving Non EMR (G)	1	1.494	6.750[1]	0.019	7.625[2]	2.002	2.795	7.981[2]	8.163[2]	15.253[2]	2.276[1]
Learning Potential (LP)	2	0.573	2.738	2.146	3.572[1]	1.018	7.403[2]	1.124	5.904[2]	8.735[2]	1.247
Sex	1	0.808	1.739	0.313	3.116	3.871	0.027	0.126	0.750	2.913	1.110
G x LP	2	0.328	0.369	1.143	0.142	0.612	1.165	0.034	0.564	0.549	0.628
G x Sex	1	0.494	0.016	6.478[1]	0.566	1.957	0.086	0.232	0.354	0.795	0.961
LP x Sex	2	1.037	2.276	0.368	1.503	1.537	0.691	0.196	0.698	1.385	0.788
G x LP x Sex	2	0.274	0.052	0.887	0.455	0.579	0.487	2.793	0.777	0.168	1.059
Error Mean Square	43	3.473	6.477	1.891	10.163	3.755	6.536	4.417	20.760	3290.720	11.298

1 $p < .05$

2 $p < .01$

differences were especially apparent on the Boards portion of the test, for the more difficult sections of the Diagrams portion (Complex Diagrams and Schematic Diagrams), and for total score. Similarly, there were differences by the learning potential main effect, most especially on subtotal (Boards), Simple and Schematic Diagrams, and total test score.

Reference to Table 20, which summarizes the F-ratios for the posttest scores, and Table 21, which summarizes the F-ratios for the gain scores, indicates that the initial differences between the groups which were evident on the pretest disappear after exposure to the electricity unit. That is, though there were differences in initial knowledge of electricity, the differences between the special and low achieving regular class children were obscured or disappeared following teaching.

The learning potential main effect indicated the only significant differences between the two groups for these posttest and improvement scores. That is, the differences in posttest score and in gain scores were largely a function of learning potential status, and not regular or special class placement. Thus high scorers and gainers tended to learn most and to have improved their scores more as a result of the curriculum unit, regardless of whether they are in special or regular class. For the posttest scores, all the subtests of the evaluation instrument were significant well beyond the one percent (1%) level, with one exception, as was the multivariate F-ratio. The results with the gain scores followed a similar pattern, six of the nine scores were highly significant ($p<.01$), and a seventh was significant at the 5% level. With few exceptions, (the learning potential x sex interaction for the Diagrams sections and total score), there were no other significant effects, on either summary table, nor for any of the other multivariate F-ratios. Figures 6 and 7 present the results comparing the two treatment groups' scores by learning potential status for the Boards and Diagrams subtotals, respectively.

The conclusion is inescapable. The hypothesis that high scorer and gainer special class children can be considered educationally rather than mentally retarded is borne out by the similarity of their learning pattern to that displayed by the dull and average IQ children who have done poorly in school. Though the special class children do know less about electricity (or display more initial discomfort with, and avoid the evaluation task on the pretest), increased familiarity with the materials and the evaluation procedure, and exposure to the curriculum unit obscures these initial differences. These findings support the hypothesis that the high scorer and gainer special class child can learn satisfactorily with presentations appropriate to their ability structure. The

Table 20. Summary of Analyses of Variance (F-ratios) of Posttest Scores for Taught EMRs and Low Achieving Regular Class Students

	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total	Multivariate F-Ratio
EMR vs. Low Achieving Regular Class (G)	1	0.799	2.002	0.025	0.241	0.894	0.804	0.863	0.436	0.391	0.563
Learning Potential (LP)	2	12.033[2]	8.190[2]	7.845[2]	19.265[2]	13.446[2]	3.623[1]	14.022[2]	14.540[2]	19.823[2]	3.092[2]
Sex	1	0.010	0.052	0.004	0.005	0.003	0.695	0.074	0.113	0.047	0.709
G x LP	2	1.559	0.094	0.252	0.376	0.471	0.478	0.240	0.251	0.268	0.746
G x Sex	1	0.554	0.337	0.006	0.000	0.001	0.487	0.061	0.041	0.019	0.310
LP x Sex	2	1.470	2.212	0.162	2.557	3.712	3.597[1]	3.019	4.350[1]	4.052[1]	1.616
G x LP x Sex	2	0.254	1.461	0.349	0.676	0.679	0.271	0.459	0.532	0.629	0.856
Error Mean Square	43	6.824	11.781	5.325	31.887	7.363	18.223	8.089	57.845	14354.445	11.403

[1] $p < .05$

[2] $p < .01$

Table 21. Summary of Analyses of Variance (F-ratios) of Gain Scores for Taught EMR and Low Achieving Regular Class Controls

	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Sub Total I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Sub Total II	Total	Multivariate F-Ratio
EMR vs. Low Achieving Regular Class (G)	1	1.980	0.215	0.004	1.472	2.633	0.009	0.791	0.928	1.056	0.965
Learning Potential (LP)	2	6.937[2]	5.416[2]	2.731	8.333[2]	9.506[2]	0.709	4.227[1]	6.008[2]	11.263[2]	2.739[2]
Sex	1	0.188	1.194	0.112	1.241	3.758	0.806	0.108	0.201	0.995	1.160
G x LP	2	0.470	0.032	0.243	0.295	0.982	0.257	0.127	0.008	0.048	0.391
G x Sex	1	0.038	0.195	1.814	0.135	1.543	0.257	0.036	0.037	0.206	0.515
LP x Sex	2	0.762	0.213	0.344	0.842	1.374	2.379	0.995	3.246[1]	2.434	0.675
G x LP x Sex	2	0.017	1.227	0.696	0.666	1.048	0.009	0.173	0.050	0.288	0.545
Error Mean Square	43	10.750	14.356	7.435	42.196	7.736	18.806	19.785	58.470	130.977	11.258

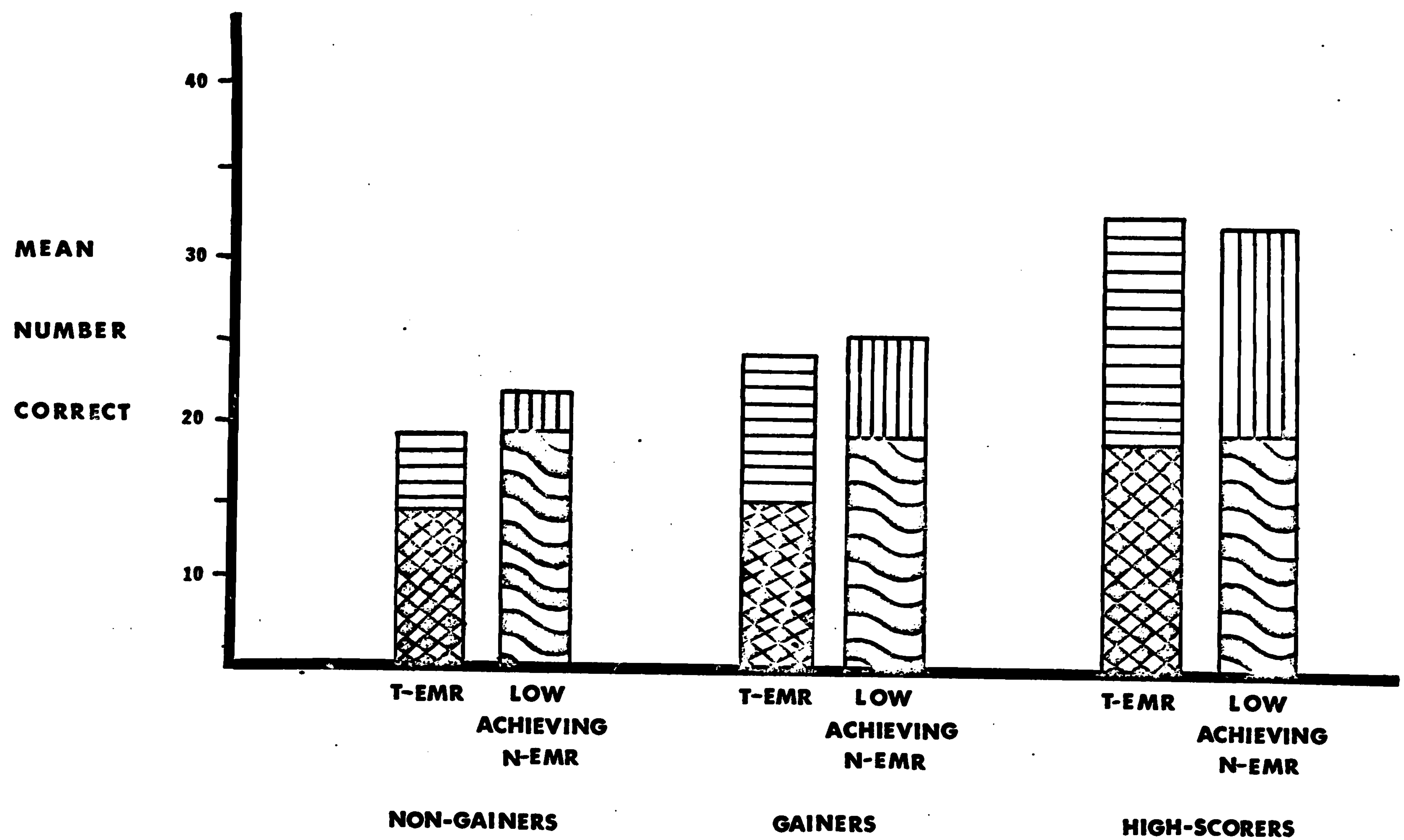


Figure 6. Comparison of Boards Subtotal Scores for Pre- and Post-test for Taught Special and Low Achieving Regular Class Students

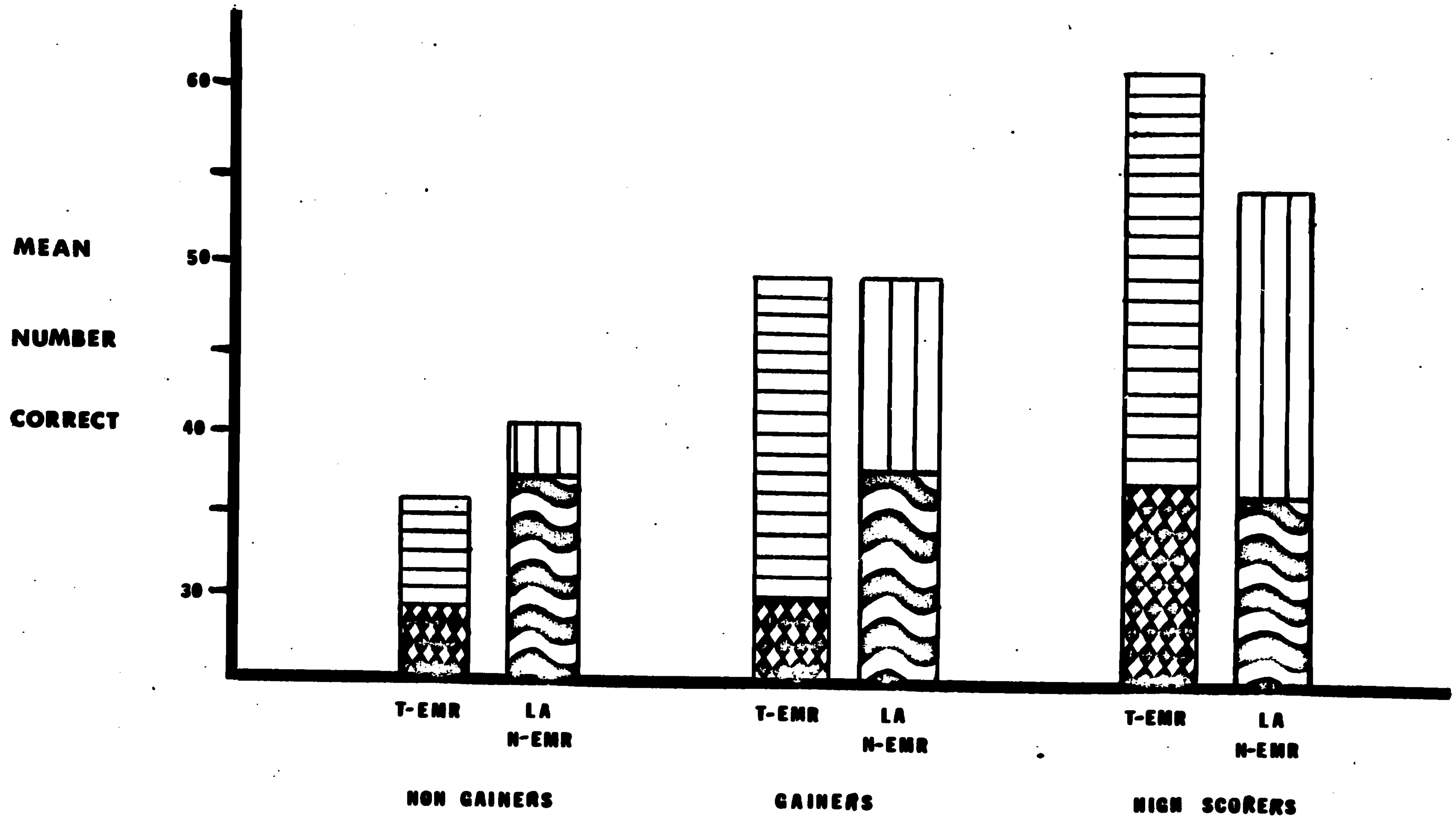


Figure 7. Comparison of Diagrams Subtotal Scores for Pre- and Post-test for Taught Special and Low Achieving Regular Class Students

evidence from this study of classroom learning suggests that their rate of learning is at least equal to that displayed by the low achieving regular age mates, at least in these special types of non-verbal learning situations.

The data from this study suggests strongly that just as low achieving regular class children with dull or average IQs may require novel presentations and interventions by which to learn that will tend to minimize the negative effects of their poor literacy skills, so too would high scorers and gainer students with IQs in the educable mentally retarded ranges seem to be able to profit equally well from similar types of specially designed classroom interventions.

B. Consistency of Response to the Electricity Concepts

1. Rationale and the Scores Explained

The differences in the ability to profit from the electricity course displayed by EMRs and regular class students of differing learning potential status was also tested by examining their ability to generalize a correct answer from one mode of presentation to another. If a subject learned the underlying concept involved in any set of matched questions, he should be able to respond to it correctly regardless of the mode of presentation. The subject who responded correctly to an item in only one presentation mode and not in the others was assumed to be responding on the basis of the specific item and not on the basis of the more general concept, demonstrating that he either had not learned the concept, or was unable to apply the concept in other instances of the same problem. The evaluation instrument was so designed that some items appeared in both the boards and the diagrams format. Some appeared twice in the diagrams section--exactly duplicating the appearance of the actual set-up, and in a schematic or symbolic format. The only difference between the items in each set was that one was presented in three-dimensional form (Boards format) and the others two dimensionally. Thus another test of differences in the learning displayed by the students in the electricity curriculum classes was their consistency of response across items in a set.

The items on the evaluation instrument were grouped into sets of identical problems and rescored for consistency of response. Twenty sets of questions involving a board and a representational diagram of the actual set-up, and 15 sets involving a board, a representational and a schematic version of the same problem appeared in the test. The sets under consideration covered 80% of the concepts taught in the course--standard brightness, bulbs and batteries appearing in parallel and series circuits, and circuits including a liquid pathway. There were too few sets for each electrical concept, hence conceptual groups were ignored and a pooled consistency score was obtained over all sets. The consistency score was obtained as follows: In a set consisting of two questions, (a boards and a representational diagram) S was given a plus (+) if he answered them both correctly, a minus (-) if he answered only one correctly and a check (✓) if he answered both incorrectly. In a set consisting of three questions (a boards item and a representational and schematic diagram), S received two pluses (++) if he answered them all correctly, one plus (+) if he answered two out of three correctly, a minus (-) if he answered only one correctly, and a check (✓) if he answered them all incorrectly. In the one set consisting of four questions, (the schematic version of the item was presented in two different ways), S was given three pluses (+++) if he answered all four correctly, two pluses (++) for three out of four correctly, one plus (+) for two out of four, a minus (-) for one out of four, and a check (✓) for answering all four incorrectly. As can be seen from this scoring system, the consistency score was dependent on being consistently right and not consistently wrong, even though when S

answered all questions in a set incorrectly he was behaving in a consistent manner. However, this latter consistency was the result of a lack of knowledge about the material involved, and not conceptual understanding. If those sets on which the students were consistently wrong were included in the consistency score, the students with the least knowledge of the material, (i.e., those who were consistently wrong) would have received the highest consistency scores.

Since it was necessary to take into account in some way those sets on which S was consistently wrong, two separate scores were obtained. In both scores the numerator was equal to the sum of the pluses. The first score was the number of pluses received divided by the total number of pluses it was theoretically possible to receive, i.e., $\frac{E(+)}{E(+) + (-) + E(\checkmark)}$. The denominator for the second consistency score was the number of times it was actually possible for the student to be consistently correct, i.e., only those items were included in which S answered at least one question in a set correctly. The two consistency scores were expressed as a percentage for each student.

The first score, which was based on the total number of times it was theoretically possible to be consistently correct, can be considered a measure of both S's knowledge of the course work and of his ability to generalize the same concept across different presentations. The greater his knowledge of the material, the greater his chances to be consistently correct. The second score excluded those sets on which S was completely wrong, i.e., where he had no knowledge of the material in question, and included only those sets on which Ss demonstrated knowledge. This score more clearly presents the generalizing inclination of the student, since it is a measure of the consistency of the student on the sets on which he displayed some understanding.

2. Results and Discussion

The two consistency scores were analyzed independently in an analysis of variance design in which the major variables were class status (taught EMR and non-EMR, or EMR taught and non-taught), learning potential status, (high scorer, gainer and nongainer), and pre- and posttest. Table 22 presents a summary of the F-ratios obtained in the four analyses of variance.

a. Consistency Score #1.

1. Taught Special and Regular Class Students.

In the first analyses, which compared taught special and regular class students, it was hypothesized that learning potential status but not class assignment would differentiate among the Ss on both consistency scores as it had in the prior analyses of the absolute

level of test scores. Thus, among the taught Ss, a taught special class child high on learning potential status would be more consistently correct than a special or regular class nongainer, but would display a level of consistency similar to the regular class students with the same learning potential status.

Regular class students were more consistently correct than EMRs, pooling over pretest and posttest scores (mean percentages were 27.20 and 18.74, for the two groups, respectively, $F=10.68, p<.01$). The main effect for learning potential status ($F=12.41, p<.001$) and the pre-posttest x learning potential interaction ($F=8.12, p<.001$) were also highly significant. As Figure 8 indicates, there were relatively small differences in the percent consistent on the pretest among the learning potential groups, disregarding class assignment, but these became quite marked on the posttest scores. The mean increment in consistency score for nongainers was only 7.3%, while the increments for gainers and high scorers were 18.8% and 22.9%, respectively. The high scorer's mean gain was three times that of the nongainers.

The pre-posttest x learning potential status x groups interaction, which was significant ($F=5.43, p<.01$), is depicted in Figure 9, and illustrates most clearly the relationship between the variables. The heterogeneity due to learning potential status within the taught groups was particularly evident. There were small differences within the special and regular class samples prior to teaching. Within the regular class sample, the effects of teaching were more determining of consistency score level. Learning potential status was a subsidiary influence, non EMRs improving less on the average (12.38%) than their gainer (20.75%) and high scorer (19.65%) peers. By contrast, changes due to differences in learning potential status following teaching were very dramatic. The special class nongainer improved minimally (3.92%), while the gainers (17.50%) and high scorers (29.11%) improved appreciably. By a consistency criterion, also, it is this nongainer group which is most likely functioning as mentally as opposed to educationally retarded. His regular class counterpart, however, starts at a higher level than any of the EMRs (including gainers and high scorers), but also improves less (12.38%) on the average than the special class gainer (17.5%) and high scorer (29.11%). Thus, although the regular class nongainer starts out at a higher level, which may be due to his richer exposure to academic materials, his ability to learn and transfer this kind of taught material was inferior to that of the special class gainer and high scorer. The achievements of these latter groups provide the most supportive evidence for the learning potential argument. Not only is there a negligible difference between the increment displayed by nonEMR and EMR gainers, (20.75% and 17.50%, respectively), but the gain between the high scorer groups favors the EMR group slightly, (29.11% and 19.65% for the EMRs and non EMRs, respectively). As can be seen in Figure 9, the EMR high scorers pretest at a lower level and posttest at a higher level than their regular class counterparts. Thus, they learned more from teaching despite the fact that they started out at an inferior level,

Table 22

Summary of Analyses of Variance (F-ratios) for Consistency Scores of Taught Special and Regular Class Samples, and Taught and Nontaught Special Class Samples.

	Taught EMR vs Non EMR			Taught vs Nontaught EMR		
	df	Score #1	Score #2	df	Score #1	Score #2
Between Ss	65			60		
Learning Potential (LP)	2	12.41^{3}	15.20^{3}	2	38.70^{3}	17.84^{3}
Groups (G)	1	10.68^{2}	10.84^{2}	1	5.30^{1}	2.51
LP X G	2	32.99^{3}	13.01^{3}	2	<1	3.18
Pre-post (PP)	1	121.49^{3}	52.09^{3}	1	40.91^{3}	93.90^{3}
PP X LP	2	8.12^{3}	3.52^{1}	2	6.38^{1}	13.77^{3}
PP X G	1	1.82	<1	1	10.41^{2}	<1
PP X LP X G	2	5.43^{2}	3.83^{1}	2	2.06	9.43^{3}
Between Ss Residual	60	146.28	459.09	55	113.05	416.94
Within Ss Residual	60	192.30	462.16	55	189.68	153.60

1 = $p < .05$
2 = $p < .01$
3 = $p < .001$

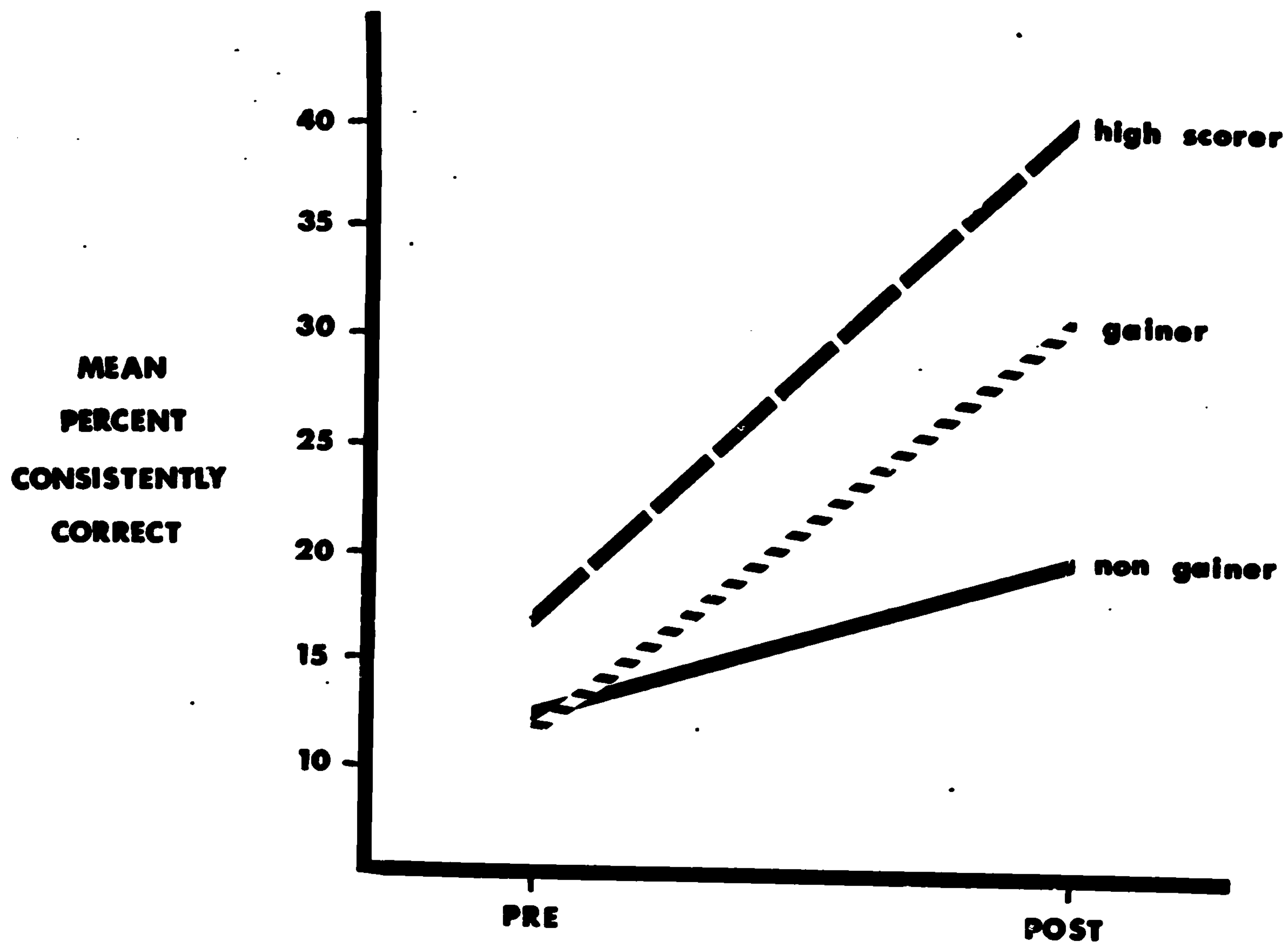


Figure 8. Taught EMR-nonEMR Comparison, Pre- Posttest x Learning Potential Status Interaction 9Disregarding Class Assignment) for Consistency Score #1.

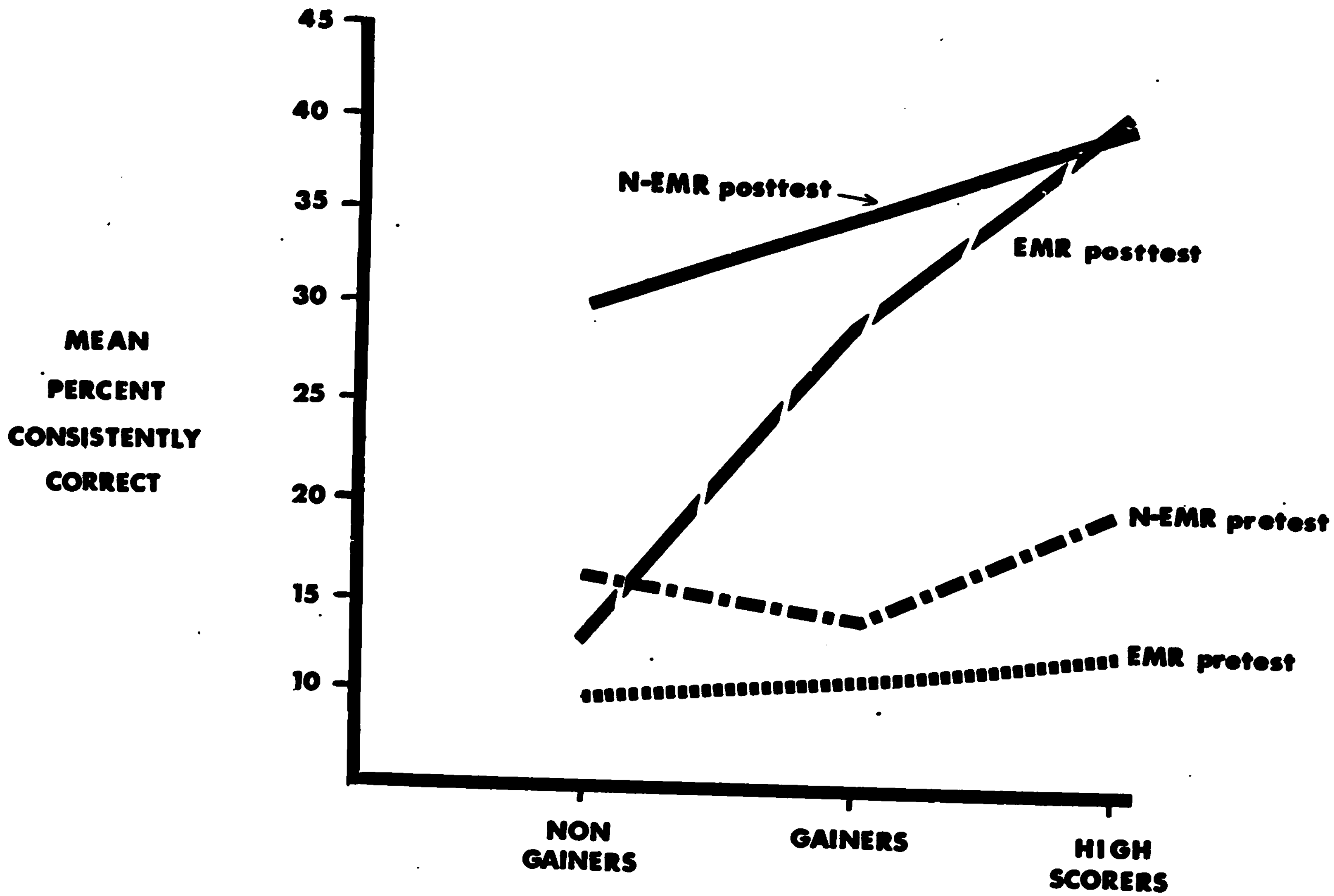


Figure 9. Taught EMR-nonEMR Comparison, Pre- Posttest x Learning Potential Status x EMR-nonEMR, for Consistency Score #1.

and were able to generalize more of what they learned, a basic factor in the learning process.

2. EMR Taught-Nontaught Comparison

Table 22, which presents the summary of the analyses of variance (F-ratios) for the EMR taught/nontaught comparison, adds additional support to the relevance of learning potential as an explanation of the heterogeneity of performance within the EMR population. There was a significant difference between the taught and nontaught groups, (main effect for teaching, $F=5.30$, $p<.05$), although both groups started from the same pretest level (10.88 and 9.96% consistent for the groups, respectively (See Figure 10)). The taught EMRs' gain was the result of having been able to acquire and retain the new material presented to them and was not just the result of a practice effect of exposure to the test.

When the Ss were categorized according to learning potential status, the main effect was significant, ($F=38.70$, $p<.001$). The scores for the three groups were not significantly different on the pretest, though the trend is in the predicted direction. However, the differences between the learning potential groups on the posttest were large and significant for both taught and nontaught subjects, (pre-posttest x learning potential interaction, $F=6.81$, $p<.05$).

Figure 11 illustrates the learning potential groups' performances, pooling over teaching condition, and indicates that even without teaching, there were differences in improvement in consistency as a function of learning potential status. The triple interaction of pre-post, learning potential status and groups, which was non-significant, illustrated in Figure 12, makes this point more clearly. The pre-posttest change for the nongainer in both taught and non-taught groups is small (mean gain was 3.92% and 3.08% for the two groups, respectively), supporting the prior findings that the non-gainer benefitted minimally from teaching even on this manipulative nonverbal curriculum. By contrast, merely the effect of repeated exposures to this evaluation instrument resulted in an improvement in the nontaught gainers' (6.13%), and high scorers' (8.33%) scores greater than that of the taught nongainers. The taught gainers' and high scorers' mean improvement was triple this amount (17.50% and 29.11%, respectively).

b. Second Consistency Score.

The second consistency score represented the proportion of pluses (+) over sets attained by a student when he answered at least one of a set correctly, disregarding those sets in which the S displayed no understanding. This measure, then, represented a purer estimate of the degree to which the S applied a principle consistently to a series of similar items in which he correctly answered

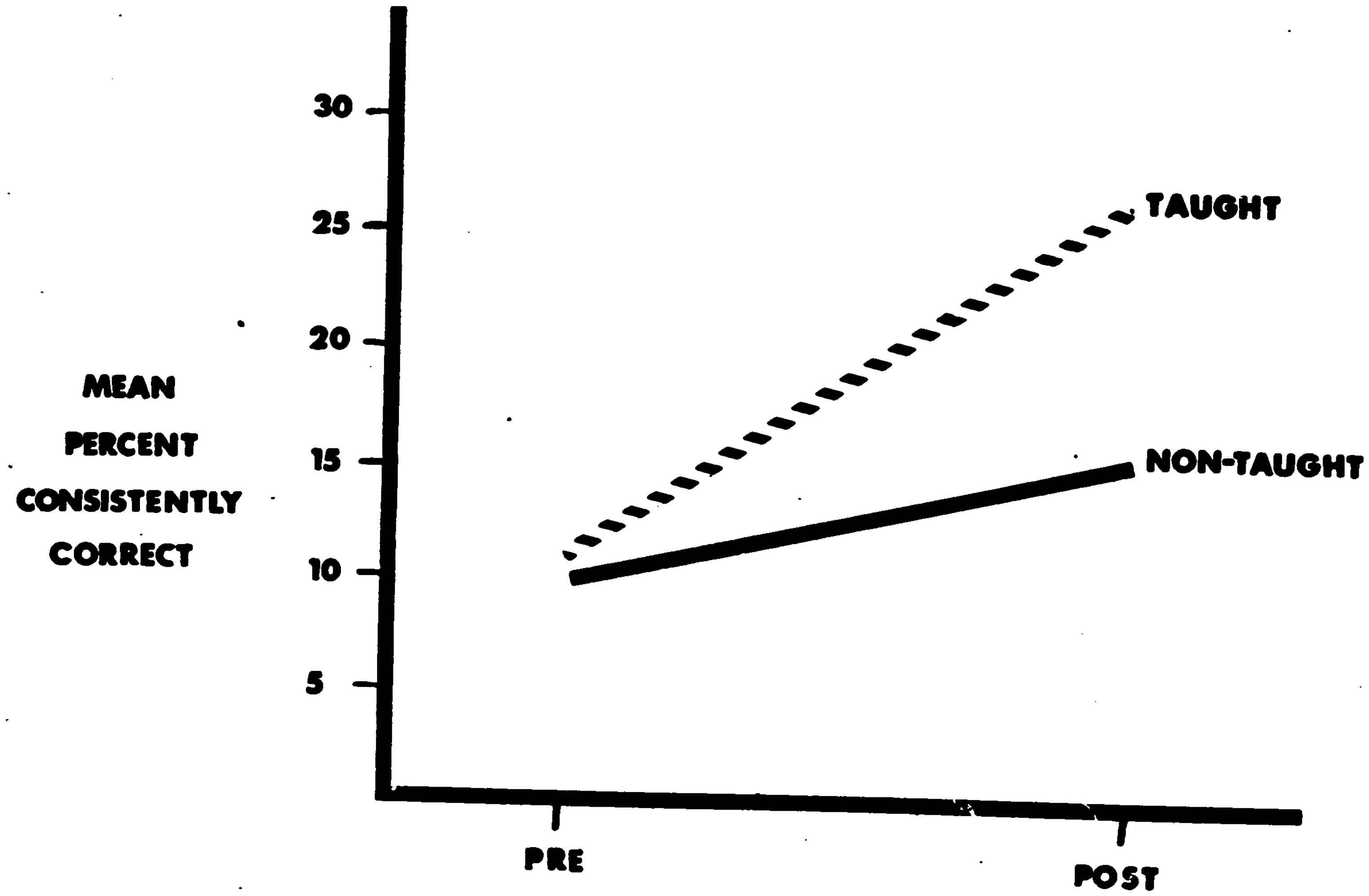


Figure 10. Taught-Nontaught EMR Comparison, Pre- Posttest x Teaching Interaction, for Consistency Score #1.

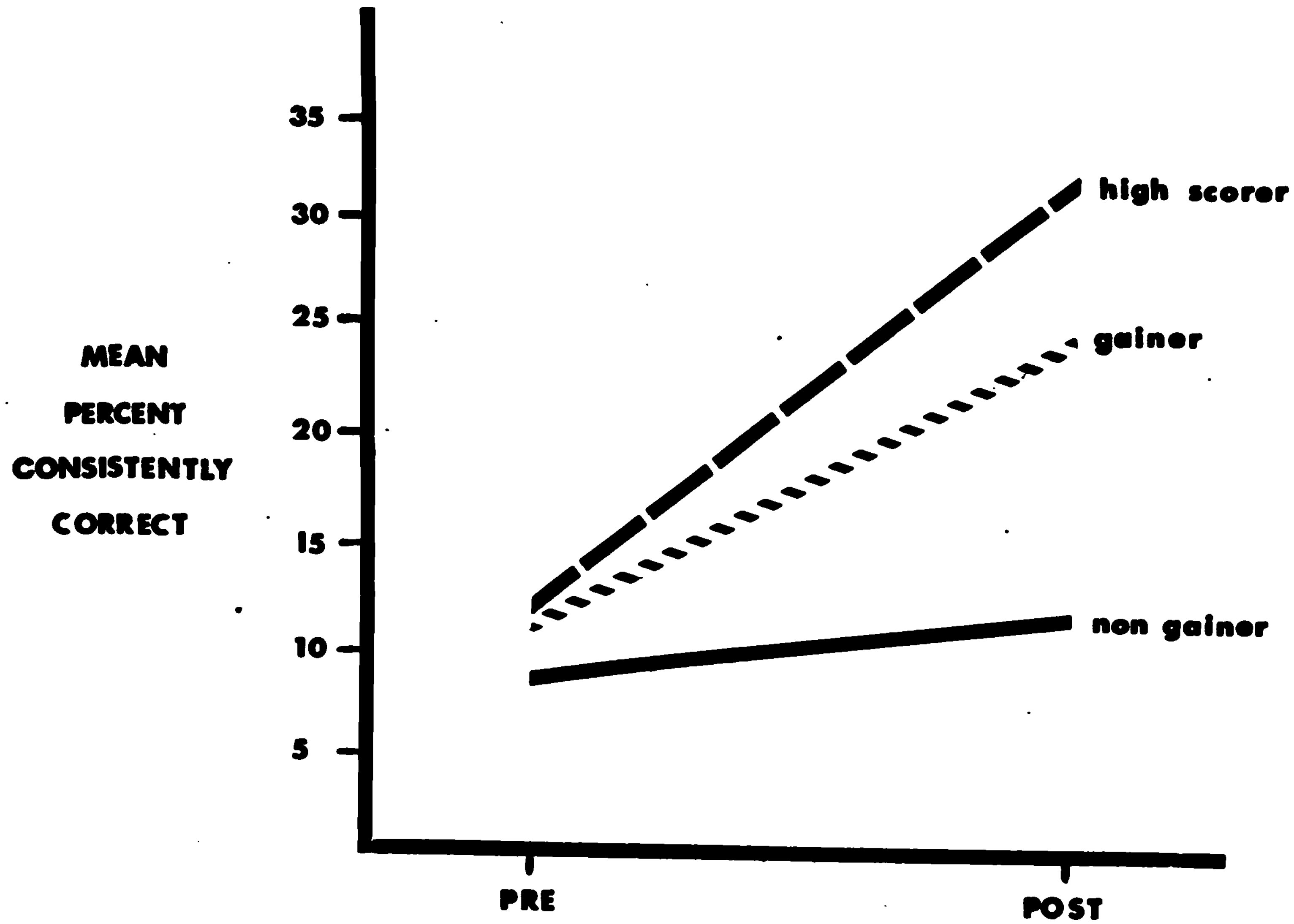


Figure 11. Taught-Nontaught EMR Comparison, Pre- Posttest x Learning Potential Interaction for Consistency Score #1.

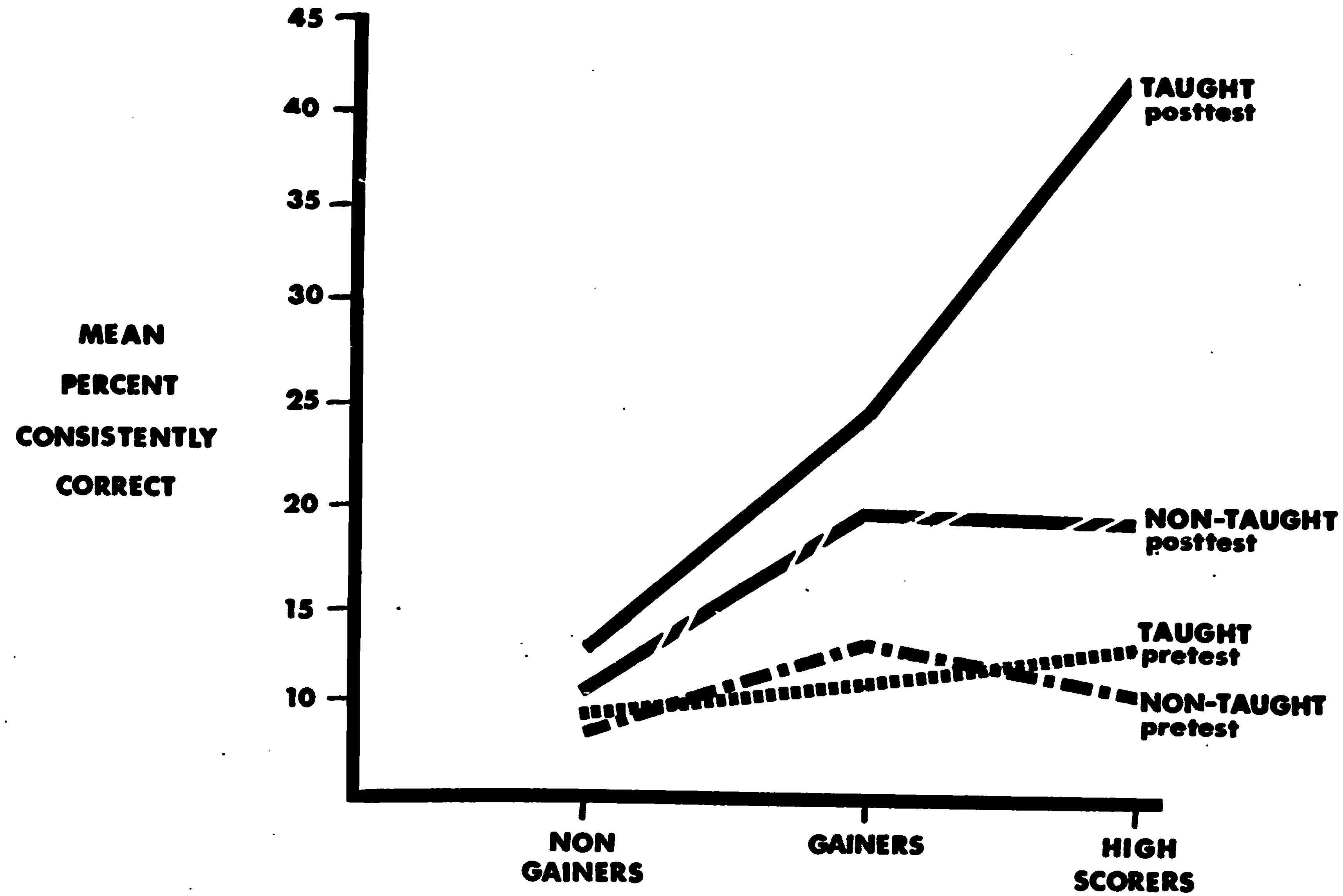


Figure 12. Taught-Nontaught EMR Comparison, Pre- Posttest x Learning Potential Status x Taught-Nontaught Interaction for Consistency Score #1.

at least one item correctly since it did not penalize him for lack of knowledge. As with the first consistency measure, two analyses of variance were run, a taught EMR/nonEMR comparison and an EMR taught/nontaught comparison, in which groups (2 levels) and learning potential status (3 levels) were the dependent measures (see Summary Table 22).

1. Taught Special and Regular Class Comparison.

As with the first analyses, it was hypothesized that differences in Ss ability to generalize across items in a set would be predicted by learning potential status, rather than special or regular class assignment. As with the first consistency score, there was a significant main effect for class assignment (F=10.84,p<.01). A major proportion of this effect however, was due to the differences between the two groups on the pretest, and to the relatively small increment in score displayed by the EMR nongainers on the posttest (see Figure 15).

As with the first consistency score, when the Ss were divided by learning potential status, there were large differences among the learning potential groups, (learning potential main effect, F=15.20,p <.001); the effect of learning potential grouping was greater within the taught special class than the taught regular class samples, (learning potential x groups interaction, F=13.01,p<.001), and there was significant pre-posttest x learning potential interaction, (F=3.52;p<.05). Figure 13 which presents learning potential status as a function of class assignment indicates very clear cut differences in scores. The nongainer, whether EMR or non EMR, falls far below the gainer and high scorer, in mean percent consistent, indicating relatively less ability to apply principles consistently to items within a set, when disregarding level of knowledge of electricity is ignored. Figure 14, which presents the pre-posttest x learning potential interaction, indicates that the initial differences among the learning potential groups, disregarding class status, were relatively small on the pretest, but were considerable following teaching.

The significant triple interaction of pre-posttest, learning potential status, and groups' (F=3.83,p<.05), more clearly indicates the interrelations among the three variables (see Figure 15). The lack of consistency on the pretest, even within sets in which the special class students knew one item, may have been due to their tendency to be "problem avoiders" when the test situation or the materials are not familiar. They tended not to apply what little information they had in a consistent manner on the pretest. Following teaching, however, gainer and high scorer special class students applied concepts equally consistently as their regular class counterparts. But, as with the first consistency score, the regular class nongainers performed more consistently before and after teaching than the special class nongainers, who improved minimally, (4.75%), even when they did answer one item in a set correctly. As with the first consistency score, Figure 15 indicates there was markedly less heterogeneity among the regular class students regardless of learning

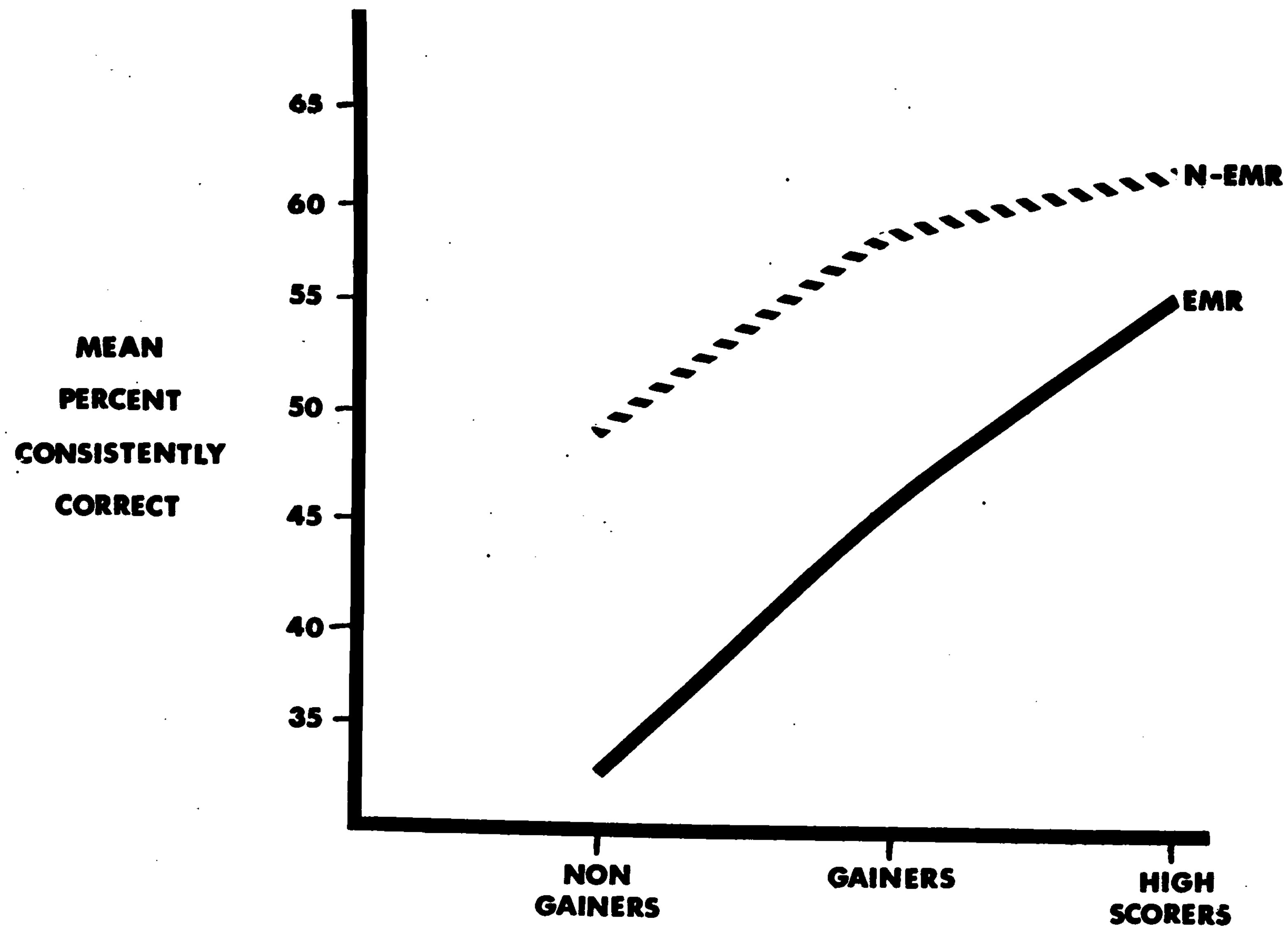


Figure 13. Taught EMR - nonEMR Comparison, Learning Potential Status x EMR-nonEMR Interaction for Consistency Score #2.

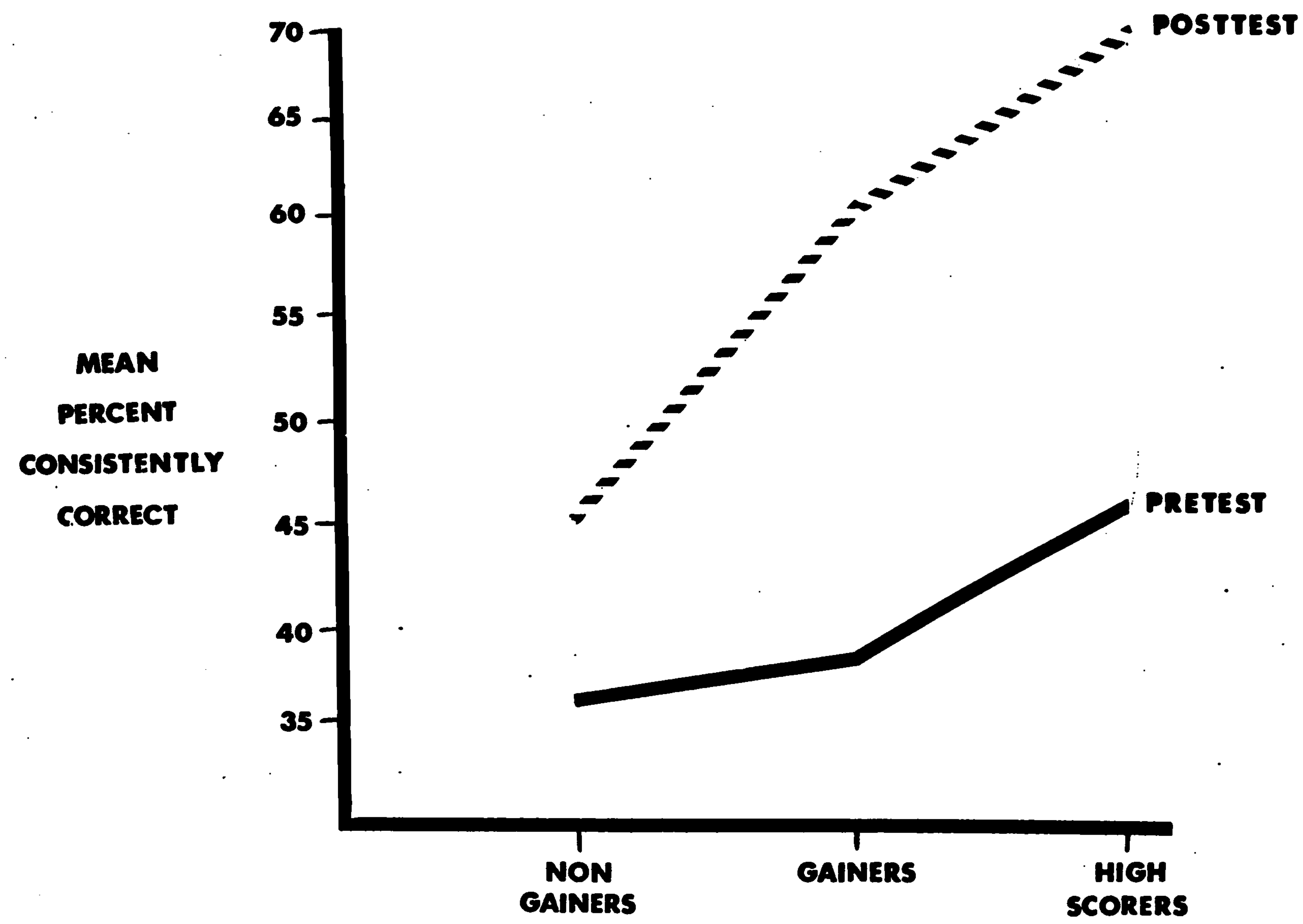


Figure 1?. Taught EMR-nonEMR Comparison, Pre- Posttest x learning potential interaction for consistency score #2.

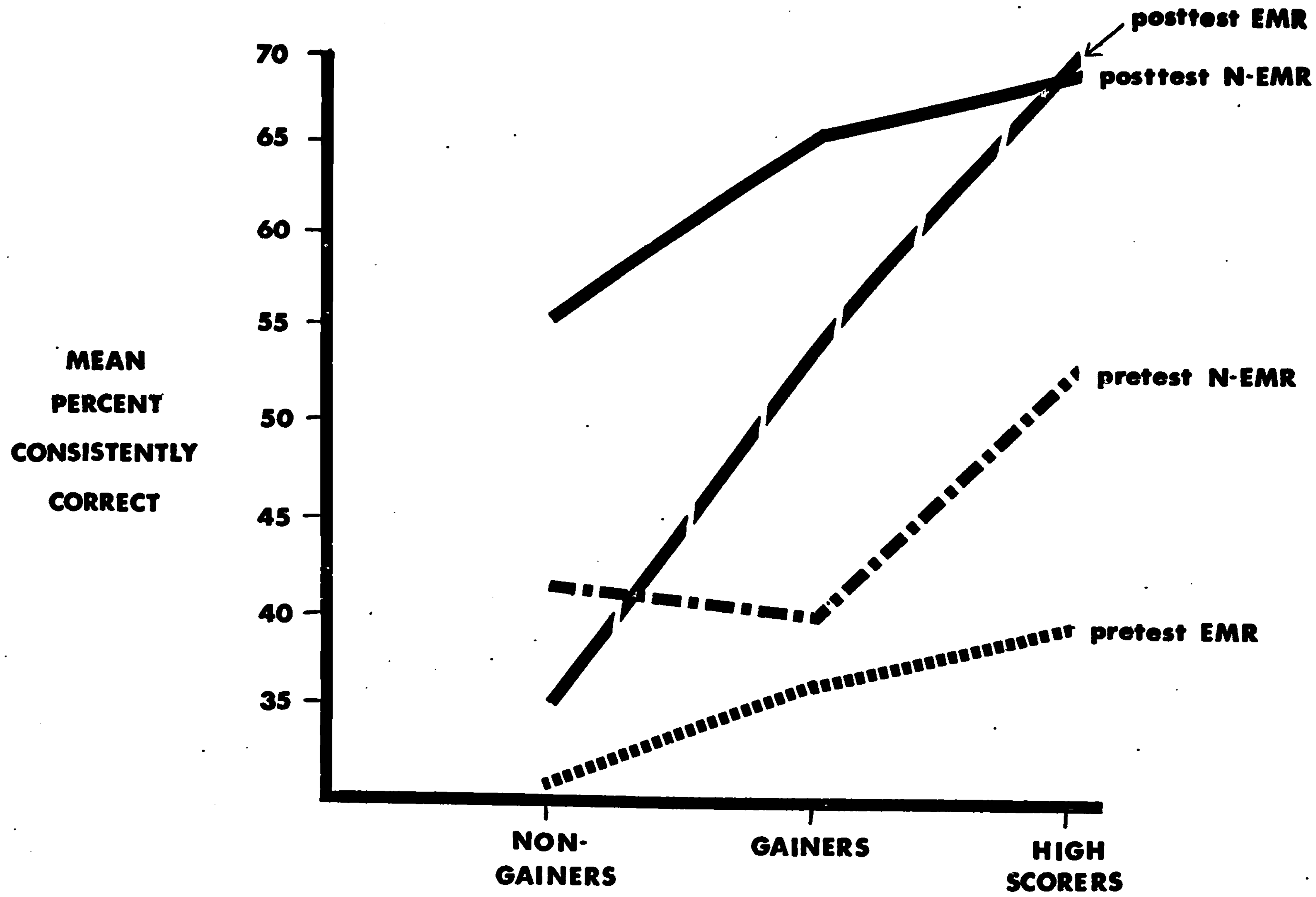


Figure 15. Taught Special and Regular Comparison, Pre- Posttest x Learning Potential Status x EMR-nonEMR Interaction for Consistency Score #2.

potential status, than among the special class students.

2. EMR Taught-Nontaught Comparison

Table 22 presents the summary of the analysis of variance for the EMR taught-nontaught comparison. The interesting aspect of this analysis is that there was no effect on degree of consistency within sets of items as a function of teaching. (Taught-nontaught main effect, $F=2.51, p<.20$); pre-posttest x taught-nontaught interaction, $F=<1$). That is, when amount of total knowledge of electricity is partialled out, both taught and nontaught special class groups displayed similar ability to apply a principle on the pre- and posttest.

However, learning potential status significantly differentiated among the students, ($F=14.58, p<.001$). While the difference between the taught and nontaught groups' means from pre- to posttest was negligible, (16.76% and 14.31%, respectively), the differences amongst the Ss in ability to generalize within sets of items was more clearly a function of learning potential status. Reference to Figure 16, (pre-posttest x learning potential interaction, $F=13.77$, $p<.001$), indicates that the high scorers' pre-posttest gain (27.34%) was almost four times that of the nongainers, (7.21%), disregarding teaching group. Figure 17, which illustrates the pre-posttest x learning potential x taught-nontaught interaction, ($F=9.43, p<.001$), indicates some small gains in percentage of consistent responses from pre- to posttest among the nongainers in both taught and nontaught groups, but more marked increments among gainers, regardless of group, and especially, among the high scorers. Nontaught high scorers did not improve their performance as markedly as their gainer counterparts. The involvement in the unit must have positively mobilized the taught high scorers' interest and involvement. The high proportion of instances in which they successfully applied principles to the item sets suggests their latent ability can be successfully mobilized in educational contexts in which they feel they can succeed. The contrast with the lower level of consistency displayed by nontaught high scorers may illustrate a motivational dimension of these students. High scorers, based on other data available, tend to be more depressed and, perhaps, more emotionally unstable than gainer EMRs (Harrison & Budoff, 1968). The initial negative anticipation may have been dissipated by the familiar posttest procedure with the nontaught gainers (who by definition are more eager to please and learn), hence, in part, their gain on the block designs task following tuition. The more emotionally unstable high scorers, who have displayed ability, may more easily express antagonism at repeated testings, and the dip in the curves on pre- and posttest may reflect motivational rather than ability factors.

In summary, then, the consistency scoring for consistent application of a concept or principle yields findings similar to the comparisons of the absolute level of evaluation test scores. Teaching the unit improved consistency of response, when knowledge of electricity (Score #1) was considered, and when the consistency score was based only on the sets in which the student demonstrated some understanding (Score #2). While regular class students tended

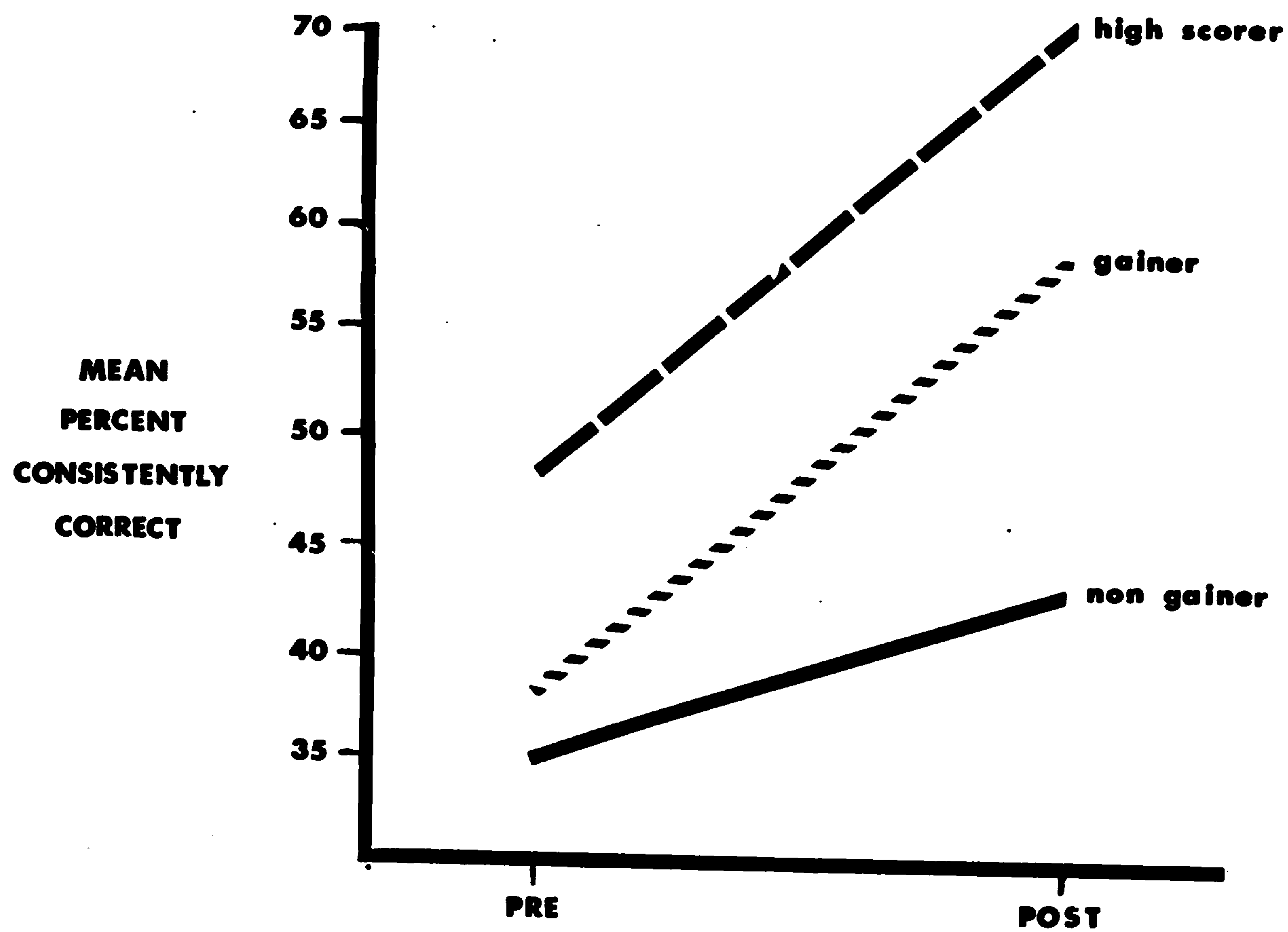


Figure 16. Taught-Nontaught EMR Comparison, Pre- Posttest x Learning Potential Status Interaction for Consistency Score #2.

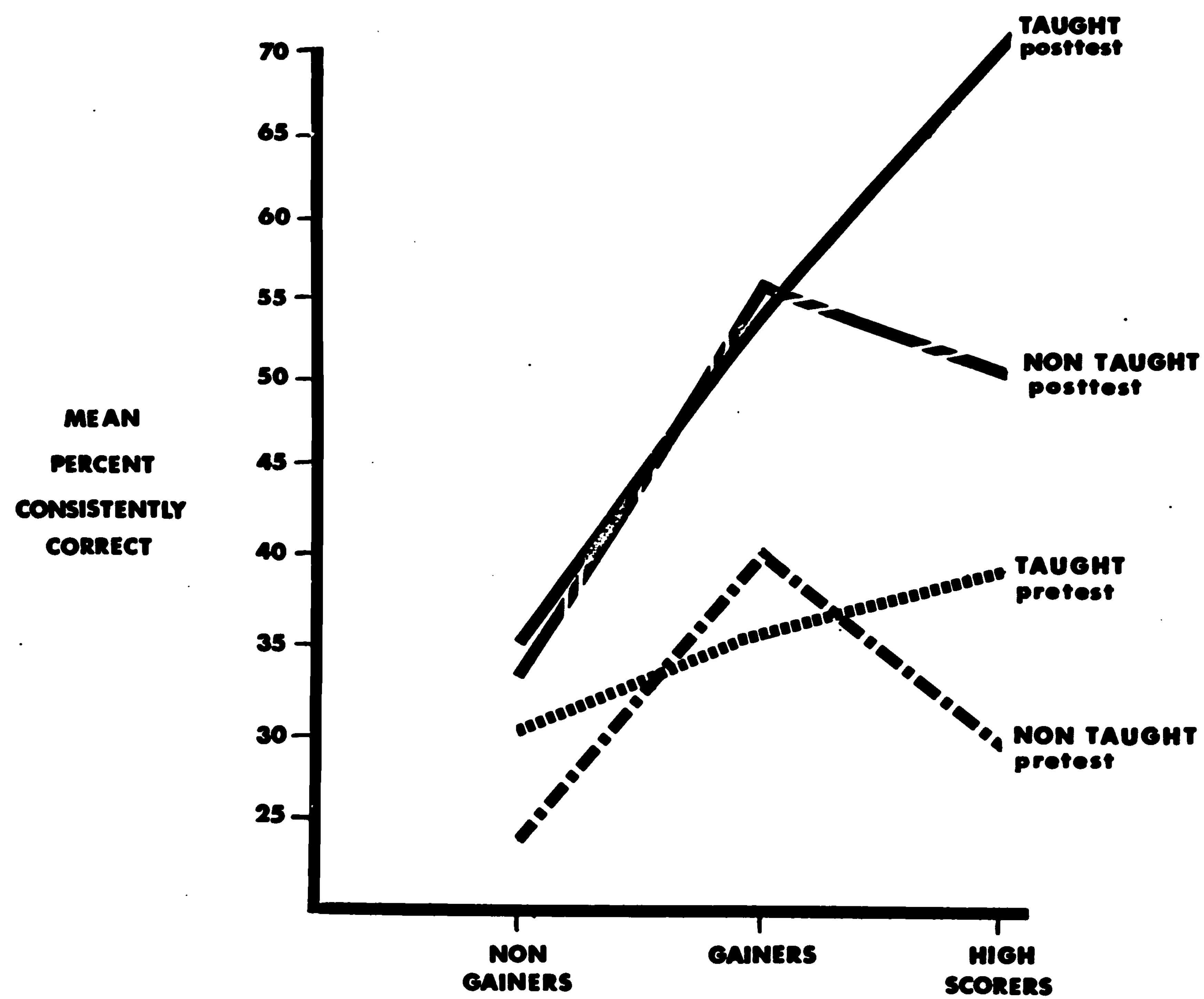


Figure 17. Taught-Nontaught EMR Comparison, Pre- Posttest x Learning Potential Status x Taught-Nontaught EMR Interaction for Consistency Score #2.

to respond more consistently on the pretest, consistent responses to the posttest was more critically a function of learning potential status. Learning potential status was the variable, singly or in combination with the pre- versus post scores, that accounted for most remaining differences in performance among special and regular class students. High scorers demonstrated the best ability, to generalize taught concepts; gainers demonstrated the next best ability, nongainers the least. But while differences in learning potential status did not result in large differences in pre- or posttest scores among regular class students, the differences were very large among the special class students, especially after teaching. Typically, even the regular class nongainer performed relatively like his gainer and high scorer peers, i.e., they tended to show appreciable gains in consistency following teaching. By contrast, the special class nongainers showed little change in consistency following teaching even when they had correctly answered one item in a set and their broader lack of knowledge was not considered (Score #2), while the special class gainers and high scorers demonstrated marked improvement, especially following teaching. Even without it being taught, gainers and high scorers improved their consistency scores somewhat. In conclusion, then, when posttest or gain scores were considered, differences between special and regular class students disappeared, and differences by learning potential status became the main differentiating variable.

C. The Relationship of Motivational, Behavioral, Social, and Psychometric Data to Scores on the Electricity Evaluation Instrument.

The grant proposal envisioned a concomitant study of motivational variables in an effort to understand more clearly what attributes the good and poor performer possessed. Various measures were proposed originally which were discarded as the project progressed.

The Parental Authority-Love Statements (PALS), a rating scale instrument of the child's perception of his mother and father, and the Projected Essential Needs (PEN-PALS), a cartoon projective technique based on eight needs and requiring a forced choice response for scoring, are parallel instruments with response categories which are comparable across the two instruments. These instruments were constructed by W. Williams (1958). The language of the questions was carefully kept within the range of third grade readers. The scales allow the child to evaluate his mother and father as he sees and reacts to them around the issues of Authority -"should or must be obeyed" - and Love - "a person who is a source of love and warmth and emotional support". Each form of the test allows a description of each parent for each of four pre-rated categories: high authority-high love, low authority-low love, low authority-high love, and high authority-low love. Algebraic summation of the responses also allows for a description of the "psychologically unknown" parent. The comparison of PALS and PEN-PALS responses demonstrated that a clinical population of delinquents tended to respond discrepantly on the two tasks, whereas the normal sample of children tended to rate their parents consistently on the two forms, providing justification for use of both approaches. Both were administered to a pilot sample of EMR subjects by reading each item aloud and having the subject choose his response. The instrument did not work with the pilot sample and was abandoned.

Sarason's recent studies of the effects of general anxiety, focalized test anxiety and defensiveness suggest the critical relationship of these variables to school performance and intelligence test scores. The General Anxiety Scale for Children (GASC), the Test Anxiety Scale for Children (TASC), the Defensiveness Scale for Children (DSC), and the Lie Scale for Children (LSC) were group-administered. Sarason, in personal conversation, had commented to the principal investigator that the EMRs to whom his group administered the DSC were remarkably high in defensiveness when compared to his normal samples. These scales were administered to pilot samples of EMRs. In fact, the correlation between anxiety and defensiveness was so high, and the mean scores so high, that the scales were deemed unsuitable since they did not discriminate among the special class students.

For various reasons germane to the total research effort, the decision to interview each participant in the study was made. Our experience with self report instruments of various types had shown them to be unreliable because it seemed the special class child did not always take them seriously enough to respond with care and understanding. In short, they seemed to yield untrustworthy data.

An individual interview which could tap a wide range of areas would provide data unavailable through self report inventories or scales. More significantly, the personal open ended contact with the student provided an opportunity for a more realistic communication, if the child was willing, or was capable of it. (A rating as to the trustworthiness of the child's response could be made as a control). The major investment then was made in this interview which is described more fully below. Only a small portion of the data is presented in this report. A fuller report will be available shortly.

The student's behavior in class and some social-demographic and psychometric data were related to the electricity evaluation scores and these correlations are presented below.

1. Correlations of Evaluation Scores with Ratings of Behavior in Class

Pearson product moment correlations were computed between the classroom observation ratings of two observers and the teachers' ratings and scores derived from the evaluation instrument. Tables 23 and 24 present these data for the special and regular class students, respectively. The observational data did not correlate with the pretest scores (nor should they) for either special or regular class students. The striking observation that can be made from the correlations between the observers' ratings and the posttest scores is that the child's motivation and attentiveness in class tend to be significantly related to posttest scores for the <u>regular</u> class but <u>not</u> for the <u>special</u> class students. This was true both for outside observers independently, and for the summed ratings of the observers.

What this suggests is that to the outside observer, the special class child's degree of involvement in the classroom tasks seemed unrelated to the learning that was taking place and that was demonstrated on the posttest evaluation. Reference to Table 23 indicates that this was not the case when the teacher predicted the children's performance on the posttest from her classroom observations. It may be that the involvement in school learning situations of the more able special class child, in a learning potential sense, may be more variable, reflecting his ambivalence toward school, although he seems to learn. The nongainer, or low learning potential child, will often be the more conforming, more diligent and persevering behaviorally, though he does not learn, as reflected by the posttest scores. The lack of significant correlations with the cooperation ratings confirms this idea. Over a sustained contact, the teacher could perceive who was making progress in understanding the material, whereas the outside observer could not. She could only rate the high learning potential child's variable behavior. What is instructive for the teacher is that when the child was given the possibility for success in understanding the subject matters, the high

Table 23

Correlational Matrix for Classroom Behaviors as Related to Evaluation Scores for Taught Special Class Students

	PRETEST			POSTTEST			GAIN		
	Subtotal		Total	Subtotal		Total	Subtotal		Total
	Boards	Diagrams	Score	Boards	Diagrams	Score	Boards	Diagrams	Score
Teacher's Ratings									
Ability	-349[1]	-208	-348[1]	-689[2]	-625[2]	-681[2]	-566[2]	-479[2]	-572[2]
Productivity & App. to Work	-106	-265	-269	-585[2]	-515[2]	-568[2]	-571[2]	-299	-452[2]
Cooperation	154	-221	-105	-254	-204	-233	-342	-016	-166
Observer 1									
Involvement	-023	-127	-121	-134	022	-049	-163	217	096
Attention	073	-160	-087	-269	-050	-151	-352[2]	118	-077
Need for Acknowledgment	-541[2]	082	-230	017	-054	-023	-262	-109	057
Observer 2									
Involvement	-023	-043	-051	-401[1]	-300	-358[1]	-427[1]	-185	-303
Attention	092	-131	-059	-309	-179	-243	-388[1]	-032	-193
Need for Acknowledgment	-204	-052	-163	-187	-155	-180	-123	-086	-099

[1] = $p < .05$.

[2] = $p < .01$.

Table 24
Correlational Matrix for Classroom Behaviors as Related to Evaluation Scores for Taught Regular Class Students

	PRETEST			POSTTEST			GAIN		
	Subtotal		Total	Subtotal		Total	Subtotal		Total
	Boards	Diagrams	Score	Boards	Diagrams	Score	Boards	Diagrams	Score
Teacher's Ratings									
Ability	-140	-317	-290	-503[2]	-510[2]	-547[2]	-405[1]	-390[1]	-444[2]
Productivity & App. to Work	086	-144	-063	-530[2]	-434[1]	-505[2]	-490[2]	-393[2]	-500[2]
Cooperation	041	-233	-147	-513[2]	-509[2]	-547[2]	-450[2]	-423[1]	-509[2]
Observer 1									
Involvement	128	-035	037	-366[1]	-440[1]	-436[1]	-350[1]	-453[2]	-476[2]
Attention	-104	-040	013	-362[1]	-380[1]	-402[1]	-341	-375[1]	-425[1]
Need for Acknowledgment	141	211	204	095	084	105	027	-016	017
Observer 2									
Involvement	043	-065	-042	-431[1]	-427[1]	-464[2]	-367[1]	-417[1]	-468[2]
Attention	008	-156	-117	-396[1]	-425[1]	-449[2]	-340	-366[1]	-415[1]
Need for Acknowledgment	134	311	275	049	014	045	007	-149	-081

[1] = $p < .05$.

[2] = $p < .01$.

learning potential student in special class learned, though his fear of failure may cause him to evidence a phenotypic behavior pattern of non-learning. By contrast, regular class students' behavior in class seemed directly related to what they learned.

2. Correlations of Evaluation Scores with Selected Social, Psychometric and Learning Data

Various social, psychometric and learning measures were available for the students participating in the experiment. The social data were derived from the child's school cumulative record as were IQ and achievement test data. Raven's Progressive Matrices (Series A,B,C,D,E) were group administered to all the students. As part of the larger research program, some of the special class students had been administered the Wechsler Intelligence Scale for Children (WISC) and a paired associate learning task. Pearson product moment correlations were computed between these various scores and the nine scores derived from the electricity pretest and posttest, and gain scores.

There were few significant correlations between the social data variables available from the child's school records and the scores obtained from the evaluation instrument. Hence, Tables 25, 26 and 27 present the correlations between selected school related and test data variables and the pretest, posttest, and gain scores on the evaluation instrument, for the taught special and regular class students, and the combined samples, respectively. Only the correlations with the Boards and Diagrams subtotals, and the total test score were included in the tables.

<u>Pretest Results</u>:

Comparing the tables, for the special and regular class students, it is apparent that the few significant correlations with pretest scores occurred mainly with performance on initial administration of the Kohs block designs for special and regular class taught subjects and latest (Binet) IQ for the special class children only. For the total sample, these measures and reading achievement level and percentile rank attained on Raven's Progressive Matrices also demonstrated significant levels of relationship.

<u>Posttest Results</u>:

For the <u>regular</u> class student, there were few significant relationships between the test variables and the scores on the electricity test obtained on the posttest. Initial level of Kohs Block Design scores (K1) and percentile rank of Raven's Progressive Matrices, both nonverbal reasoning tasks, were the only measures which correlated significantly with posttest performance. By contrast, for the <u>special</u> class child, scores on pretest (K1) immediate (K2) and delayed (K3) posttests on the Kohs block designs, and Raven's Progressive Matrices, were significantly correlated (most at

Table 25. The Correlation of School and Test Variables to Pretest, Posttest and Gain Scores on the Electricity Evaluation Instrument for the Taught Special Class Students

	PRETEST			POSTTEST			GAIN		
	Subtotal Boards	Subtotal Diagrams	Total Score	Subtotal Boards	Subtotal Diagrams	Total Score	Subtotal Boards	Subtotal Diagrams	Total Score
Sex	-434	-325	-487	-265	-316	-307	-064	-111	-085
Age - 1st in Special Class	209	291	331	141	412[1]	297	038	335	264
No. Years in Special Class	020	-249	-187	199	-092	045	227	-047	041
Pretest, block designs (K_1)	274	324	386[1]	575[2]	510[2]	559[2]	506[2]	329	448[2]
Immediate Posttest (K_2) (N = 25)	283	-048	146	626[2]	416[1]	552[2]	466[1]	304	442[1]
Delayed Posttest (K_3) (N = 23)	276	207	339	747[2]	551[2]	682[2]	573[2]	347	535[2]
Latest Binet IQ (N = 20)	264	488[2]	522[2]	340	370[1]	363[1]	232	128	204
Verbal Scale IQ (N = 11)	389	298	430	126	373	283	-243	333	216
Performance Scale IQ(N = 11)	165	435	413	519	354	428	421	213	446
Reading Achievement	068	-043	005	-001	-044	-030	-045	046	033
Math Achievement (N = 17)	482	199	392	386	516[1]	475	228	443	370
Raven's Percentile (N = 23)	-033	288	204	716[2]	576[2]	666[2]	753[2]	423[1]	649[2]
PA Forwards Errors (N = 8)	-329	-120	-212	-002	141	068	070	256	148
PA Backwards Errors (N = 7)	-756[1]	-375	-586	-335	-335	-348	-173	030	-126
Gain K_2 - K_1	026	163	143	106	-028	028	105	-103	-011
Gain K_3 - K_1	181	308	330	179	414[1]	316	094	325	283
Learning Potential Status	147	284	286	223	310	278	171	231	281

Except where otherwise indicated, N = 33.

[1] = $p<.05$

[2] = $p<.01$

Table 26. The Correlation of School and Test Variables to Pretest, Posttest and Gain Scores on the Electricity Evaluation Instrument for the Taught Regular Class Students

	PRETEST			POSTTEST			GAIN		
	Subtotal		Total	Subtotal		Total	Subtotal		Total
	Boards	Diagrams	Score	Boards	Diagrams	Score	Boards	Diagrams	Score
Sex	-158	-083	-144	-089	-085	-094	021	-055	-040
Pretest, Block designs (K_1)	278	479[2]	458[2]	417[1]	328	382[1]	224	132	209
Immediate Post-test (K_2) (N = 25)	-062	185	099	389	289	343	419	224	306
Delayed Posttest (K_3) (N = 23)	-061	241	146	240	155	196	259	062	137
Latest Binet IQ	-157	064	-040	255	293	285	301	274	320
Reading Achievement (N = 51)	233	271	276	046	046	046	016	-077	-079
Raven's Percentile (N = 25)	-231	000	-121	532[2]	520[2]	550[2]	660[2]	538[2]	634[2]
Gain $K_2 - K_1$	-291	012	-136	281	161	231	246	184	307
Learning Potential Status									

Except where otherwise indicated, N = 33.

[1] = $p<.05$.

[2] = $p<.01$.

Table 27

The Correlation of School and Test Variables to Pretest, Posttest and Gain Scores on the Electricity Evaluation Instrument for the Total of Taught Students

	PRETEST			POSTTEST			GAIN		
	Subtotal		Total	Subtotal		Total	Subtotal		Total
	Boards	Diagrams	Score	Boards	Diagrams	Score	Boards	Diagrams	Score
Sex	-274[1]	-182	-266[1]	-170	-187	-190	-021	-081	-058
Age - 1st in Special Class (N = 33)	120	225	228	089	373[1]	253	032	325	250
No. Years in Special Class (N = 33)	088	-190	-100	232	-067	072	224	-047	044
Pretest, block designs (K_1)	362[2]	479[2]	512[2]	537[2]	476[2]	522[2]	359[2]	249[1]	346[2]
Immediate Posttest (K_2) (N = 41)	270	110	205	561[2]	389[1]	484[2]	449[2]	283	372[1]
Delayed Posttest (K_3) (N = 39)	288	304	356[1]	594[2]	438[2]	523[2]	470[2]	261	378[1]
Latest Binet IQ	284[1]	465[2]	475[2]	357[2]	408[2]	396[2]	203	199	255[1]
Verbal Scale IQ (N = 11)	389	298	430	126	373	283	-243	333	216
Performance Scale IQ (N = 11)	165	435	413	519	354	428	421	213	446
Reading Achievement (N = 51)	291[1]	322[1]	375[2]	159	160	164	016	042	059
Math Achievement (N = 17)	482	199	392	386	516[1]	475	228	443	370
Raven's Percentile (N = 48)	090	388[2]	321[1]	695[2]	637[2]	687[2]	682[2]	467[2]	635[2]
Gain $K_2 - K_1$	-113	066	-008	145	021	078	154	001	104
Gain $K_3 - K_1$ (N = 34)	098	245	232	129	378[1]	274	087	316	270

Except where otherwise indicated, N = 66.

[1] = $p<.05$

[2] = $p<.01$

p<.01) with the posttest scores derived from the evaluation instrument. Appreciable correlations for this group, which were nonsignificant because of the small number of scores available, were evident between the posttest evaluation scores and arithmetic achievement, Wechsler performance scale IQ, and a paired associates learning measure.* The "backwards" learning task seems more closely related to posttest scores on the evaluation instruments than the usual paired associates learning paradigm (PA Forwards). The explanation for this finding is that fewer errors on the second session after an interval of several days indicates the degree to which learning had taken place, as opposed to initial grasp of and memory for the pairs which may be more critical determinants of the usual paired associate learning. What is interesting is that though the latest Binet IQ was highly related to pretest score, it was not related to posttest or to gain scores on the evaluation instrument. In a similar vein, Wechsler verbal scale IQ and reading achievement were uncorrelated with posttest gain scores among the special class students and for the total sample.

Gain or Improvement Scores:

Gain or improvement scores (posttest minus pretest scores) showed the same general trends. The few correlations of any magnitude appear with special class students and tend to appear with the nonverbal measures (pretest (K1), immediate (K2) and delayed posttest (K3) on the Kohs block designs, Wechsler performance scale IQ, Raven's Progressive Matrices), but not with verbally dependent measures such as Binet IQ or reading achievement.

These correlational findings tend to support the argument and findings of the study. For the special class students, especially, scores on the posttest evaluation correlated more closely with the nonverbal reasoning or learning tasks and with arithmetic achievement than with pre-teaching scores. That is, the valid measure of their ability to profit from experience is represented by the measures which tap reasoning and learning abilities which are not dependent on the verbal-expressive or conceptual skills which heavily weight the scores derived from the usual intelligence tests. The appreciable degree of relationship between the knowledge attained and applied in the evaluation test following instruction and the nonverbal reasoning and learning scores adds further substance to this argument, in that these nonverbal measures do tap ability which is not reflected in the usual academic areas, e.g., reading achievement, or in

*This task was presented in two formats, the usual stimulus-response presentation (forwards paired associates learning), and at a second session in which the children were required to pair the picture earlier presented as the response member to the earlier presented stimulus (backwards paired associate learning).

measures of scholastic attitude, e.g. Wechsler Verbal Scale IQ or (and) Stanford Binet IQ. The low correlations with reading achievement provide substantive support that the electricity course and the evaluation unit did not require reading or verbal expressive skills for a successful performance.

3. Relationship Between Interview Variables and Electricity Evaluation Scores.

The taught special and regular class students and the non-taught special class students at the school in which the classes were taught were interviewed for one hour as an additional means by which to explore motivational parameters which relate to the heterogeneity of performance displayed by the special class sample.

An interview was developed to cover the three main areas of the student's life: social, family and academic. Although the situations with which the students were presented were different in each section, depending on the area involved, the interview was designed so that the phrasing of the questions and the scoring systems were comparable across sections. In this way discrepancy scores could be obtained between subjects and also between sections of the interview for a given student.

Interview Data

Three main types of questions were given in each section of the interview, sociometric, role playing, and striving behavior questions. On the sociometric questions S was presented with a situation (e.g., social - want to go to a movie; family - unhappy and need someone with whom to talk; and academic - need help with homework) and was asked for the persons he would and would not choose for that specific situation. He was required to give the name of a child and an adult. The object of these questions was to see with whom S identified, whether or not his peers were from special or regular classes, and whether they changed as the situation changed.

S was also again given a situation into which he was supposed to project himself for the role playing questions, (club activities, household duties, class participation). He was given a choice of four roles that were available (ranging from observer to leader) and was asked to name the one for which others would choose him and the one he would like to have. The point was to see whether or not S's role outside of school was the same or different from his classroom role.

For the striving behavior questions, S was asked to state the way things are and the way he would like them to be. He was asked whenever possible (on a prearranged schedule) how he would like things to be if he could have his way. Then, whenever a discrepancy occurred, S was asked if things could be changed, and if so, how. The aim was to see whether or not he sees himself as having any control over situations and if so, how and in what areas. It was hoped that from these questions a measure of Ss striving behavior would be obtained.

Table 28. Correlations Between Electricity Evaluation Scores (Pre-, Posttest and Gain) and Selected Interview Variables for the Taught EMR Sample (N=31)

	PRETEST			POSTTEST			GAIN		
	Subtotal		Total	Subtotal		Total	Subtotal		Total
	Boards	Diagrams	Score	Boards	Diagrams	Score	Boards	Diagrams	Score
School Value Locus	-059	.523[2]	.373[1]	344[1]	219	275	390[1]	-156	063
Total School Effort Score	044	187	158	456[2]	217	332	484[2]	052	244
Total Social Responsibility	031	.220	.189	415[1]	202	312	436[1]	012	194
Total Social Role	131	300	293	421[1]	296	358	388	115	260
Real Academic Responsibility	075	015	035	263	217	248	265	239	283
Ideal Future Job Score	248	549[2]	557[2]	462[2]	500[2]	494[2]	351[1]	248	345[1]
Family Role Score I	-020	-116	-100	-400[1]	-251	-325	-413[1]	-113	-251

[1] = $p < .05$
[2] = $p < .01$

Each of these types of questions, sociometric, role playing, and striving was presented at least two times per section so there would be some measure of consistency of response.

In addition to these three types of questions, each section had questions specific to the problems being investigated in that area. In the social section, assuming the EMR's interests are not in school, the attempt was made to see whether or not S had any outside interests, and if so, what they were and how involved he was in them. With this in mind S was given a series of questions related to leisure activities, clubs, lessons, and after school employment. Each short series tapped his degree of involvement, likes and dislikes, reasons for wanting or not wanting to engage in the activities, etc. Each question and answer had a weighted score assigned to it so that at the end of a section a picture could be obtained not only of each individual activity but also of S's total actual and desired outside interests.

This same type of scoring system was applied in the family section. The attempt was made to obtain a picture of the kinds and amounts of interaction that S would or would not like to have with his family. To obtain this overall score, S was presented with different situations (i.e., social events, everyday matters in the home, important family decisions, typical family problems, etc.) and was asked, depending on the question, whom he would choose to involve in the situation, which situation he would or would not involve himself in, as well as to state to whom he thinks his parents would choose to turn. In each instance S first reported what actually existed in his home, and then was asked how he would like things to be if he could have his way. In this manner, some picture of the amount of satisfaction or dissatisfaction that exists for S in his home life was obtained.

The purpose of the academic section was to find out how the special class student, as compared to the regular class student, sees himself academically in relation to his classmates, siblings, etc., what his aspirations are regarding future schooling and career, the importance or unimportance of going to school and how it affects his future; also his likes and dislikes about his courses. And as in the other sections, S was also asked how he would like things to be for him in school if he had his way. The questions in the section were more direct than in the previous one. It was only on the Locus of Control and the responsibility questions that S was presented with situations into which he was asked to project himself. The Locus of Control scale consisted of a series of twelve questions on success and failure situations in the school setting. There were an equal number of success and failure questions. For each type of success situation that S was given, he was also presented with an identical failure situation. The purpose of the Locus of Control questions was to measure the amount of responsibility S assumed for his successes and failures in school as opposed to projecting them onto his teachers, parents, etc.

The responsibility section was the last part of the interview. It consisted of a series of a dozen questions, each one based on a situation of conflict (i.e., go with friends vs. finish job for parent). There was an equal number of questions devoted to each of the three areas, social, family, academic. Each type of conflict represented in one area is presented exactly the same way in the other areas with a change in the people involved i.e., friends, parents, teachers. In this way, as was done throughout the interview, the results between areas can be compared as well as summed across areas to obtain a total responsibility score.

The aim of the interview at all times was to compare the three learning potential groups within the EMR sample with each other, and the EMRs as a group with the non EMRs.

The variable scores selected for comparison with the evaluation scores from the curriculum study were those which related to interest in school, responsibility, and academic aspirations, and relevance to future job. It was assumed that those Ss who were more intrinsically motivated in regard to school, were more responsible academically and socially, had more realistic and higher academic or job aspirations, etc., would perform better on the evaluation instrument. It may be that the low scholastic aptitude, represented by the low IQ scores, could be overcome by salutary motivations, especially in the context of a course that did not force the students to rely on their poor reading skills, and hence maximized their possibility of success in learning the materials.

Correlations were run for each of the three groups (EMR taught and nontaught, and taught nonEMR) with twenty-eight of the interview variables and the pre- and post subtests and total test scores on the electricity evaluation instrument. Contrary to what was expected there was very little relationship between the personality measures and the test scores for the regular class sample. The special class taught sample had more significant correlations which tended to be in the predicted direction. The special class nontaught group behaved in a comparable way to their taught classmates on the pretest correlations but in an opposite manner on the posttest. The latter group consisted of only the eleven students in the same school who had not been included in the electricity classes. Because of the small number of students, these correlations are not reported in their entirety. Some suggestive relationships are reported because they support or contrast interestingly with those of the taught special class sample.

Since the correlations are the most numerous and most consistent for the special class sample, only these are presented in Table 28. There was a positive relationship between the "School Value Locus" score and five of the scores derived from the evaluation tests, including scores on pre- and posttests. These significant correlations indicate that those Ss who reported themselves as intrinsically motivated in regard to their school work obtained higher test scores on the pretest and on the posttest. Consistent with this finding was the positive relationship between total "School Effort" score and test results.

That is, those students who saw themselves as working hard and trying their best do attain higher scores. These two relationships suggest the more "successful" special class student has not yet been completely defeated by his lack of progress in school. He still sees some purpose in going to school and has enough of a sense of responsibility to work towards whatever benefits education can give him. Furthermore, his "Choice of his Ideal Future Job" relates strongly to his pre-test posttest, and gain scores-- another indication of strong involvement rather than alienation. A high score on this variable indicated that S reported obtaining his ideas about his future career by either actually participating in the work or by talking to the persons who did this kind of work as opposed to having just thought up the idea by himself without any reality testing. That high scoring EMRs possessed a greater sense of responsibility for their future and were involved in it was given further support by the positive correlations of test scores and such variables as Total Social Responsibility, Total Social Role score, Total Role score and Academic Responsibility (real). What was interesting about these latter relationships was that they occurred primarily with the posttest scores. That is, the expressed higher responsibility was most apparently related to the level of performance the special class student attained following exposure to the laboratory science unit, measures that depended in part on involvement and sustained effort, rather than on the level of performance prior to involvement.

Those students who were most able to benefit from the electricity classes were also those who reported that they are chosen for more responsible roles by their peers in social situations (Total Social Role and Total Role scores) and by their teachers in academic situations (academic responsibility (real)). In line with their more responsible social role, they also attained a high score on the social responsibility scale, indicating that they are well aware of their responsibility to others in social situations. Thus, among special class students, the various self report measures related to school performance are best validated by high correlations with performance following an opportunity in which to obtain a sense of competence with the subject matter, rather than with initial level of performance on the electricity evaluation, which provides additional support for the learning potential assessment paradigm.

The negative correlations between Family Role Score (I) and the test scores occurred because the low scores on family responsibility were defined as stated responsibility for himself and his own belongings, rather than a more general involvement in family duties.

Nontaught Special Class Students.

The sample of nontaught EMRs in the same school was small (N=11), and these correlations are not reported. Two variables (Total Effort Score and School Value--Locus of Control) supported the results with the EMR taught sample, in that there were positive correlations with pretest scores. Thus, students who reported themselves as working ard and intrinsically motivated for school tended to do well on the pretest. There were no correlations with posttest scores because whatever involvement was present for the pretest was not sustained for the posttest, as would be expected for this nontaught sample, who saw their fellow students going to classes but were excluded from them.

Regular Class Sample.

There were few significant correlations found between interview variables and test scores in the regular class group. Contrary to the pattern evident with the taught special class sample, the few significant relationships tended to occur with the pretest scores. It may be the previous history of successful learning in school, and outside it means that present level of functioning, as indicated on the pretest, more than ability to improve, may summarily relate better to responses in other life areas for the regular class student. But the stigmatized special class students' hopes, which are reflected in their interview responses, may be more accurately related to "what they might do if....". This is the situation which came following a period in which they were given the opportunity to learn. If these relationships between special and regular class students can be validated with other motivational variables, these findings would offer additional support for the learning potential type of assessment procedure with special class youngsters.

V. DISCUSSION

The major hypotheses of this study were supported. The teaching intervention did successfully increase the taught students' knowledge of the simple concepts of electricity with a unit that was an academic subject, and was taught in a highly motivating format, without the need for reading, verbal expression, or listening to teachers' lectures. In this special curriculum the student had the opportunity to concretely manipulate materials such that he could learn the concepts himself as he constructed the appropriate setups, and tested their characteristics. In this type of educational context, it was hypothesized that high learning potential special class students (high scorers and gainers) from socio-economically poor backgrounds would perform more poorly than adequately achieving CA peers of dull and average IQ from similarly poor homes. This hypothesis was only partially supported in that while the CA controls demonstrated a clearly superior level of performance on the pretest, their performance on the posttest was not as clearly superior (the F-ratios were smaller on fewer subtest scores), and there was no clearly evident superiority of performance when the gain scores were analyzed. That is, the special class students improved their performance from pre- to posttest to the same extent, in magnitude of score, as the CA controls.

Equality of performance was hypothesized between the high learning potential special class and the low achieving regular class students and this parity of performance was evident. Again, there was an initial superiority of performance on the pretest which was no longer evident on the posttest or the gain score comparisons.

The same results were obtained in the analysis of the students' consistency of response to sets of items in which the same concept was presented. The regular class students obtained higher consistency scores, by both criteria, on the pretest but this disparity in performance was not as evident on the posttest.

It must be noted, however, that the CA control sample was composed largely of low achieving students, i.e., two thirds of the sample attained grade point averages below a C (2.0) and their mean reading achievement score was more than two grades behind their grade placement.* That the CA controls were not

*The mean reading achievement scores of the total regular class sample by learning potential status was a high scorer, 5.01 (± 1.07), gainer, 5.75 (± .62) and nongainer, 4.08 (±1.00). By contrast, the mean reading achievement score of the taught special class sample by learning status was 3.42(± .72), 3.52 (± .60) and 3.36 (± .78), for high scorers, gainers and nongainers respectively.

clearly superior in their posttest performance may have been a function of the fact that so many were low achievers. It cannot be concluded then, that the special class sample did perform no differently than the CA adequate achieving controls. One may say that they did perform as well as low achieving CA regular class students who also have been experiencing considerable difficulty performing satisfactorily in school, but there were too few CA adequate achievers in the sample to properly test the first hypothesis.

One explanation for the usually low scores of the special class child on the first test administration is that he initially responds avoidantly to strange and unfamiliar situations that are evaluative. This prevents him from effectively applying information or concepts he has or learns during the test to subsequent items. This failure to relax enough to use such clues in test taking depresses his final score, though it probably does not handicap the regular class child as much, especially one that has been successfully achieving in school. This is one reason why the block design learning potential procedure relies on multiple exposures to familiarize the child with the problems before final assessment of his ability. If this factor accounts for some of the pretest differences on the electricity evaluation instrument, then the special class student overcomes this initial handicap since his posttest scores were at the same absolute level as those of the low achieving regular class child. The evidence of this study indicates the special class child learns at a similar rate to the low achieving regular class child when a nonverbal manipulative curriculum is used in which he can become conversant with the concepts.

The most significant findings of the study relate to learning potential status. Though this factor tends not to be a significant determinant of scores on the pretest, when regular and special class students were compared, it is a strongly significant determinant of absolute score level, and in the degree to which the students in both groups consistently apply a concept within both groups following teaching. In the comparisons of the scores attained by the low achieving regular and special class students, it was the only between groups variable which was significant in the posttest and gain analyses. It was only this factor which tended to account for the differences in score among individuals within the two samples following teaching. In all comparisons, as reference to the tables and figures indicate, high scorers tended to improve their understanding more than gainers, and nongainers improved least, regardless of whether they were in a special or regular class.

It was among the special class sample that the significance of the learning potential distinctions were most dramatically apparent. Special class nongainers essentially gained little or nothing in their understanding of the electricity concepts, or in their ability to generalize these concepts even when their limited knowledge was not considered following teaching. (Consistency Score #2). By contrast, the high scorers' performance

was often superior to that of their regular class peers as their post-teaching scores were often highest of all the subgroups. The gainers' attainments were also considerable. There was relatively less difference among the attainments of the regular class students by learning potential status. Thus, though regular class nongainers gained less on the average than their high scorer or gainer classmates, all three groups tended to improve their score appreciably following teaching. The differences among the regular class learning potential groups, however, were of sufficient magnitude to contribute to the learning potential effect across special and regular class groups.

Within the comparisons of taught and nontaught special class students, while the taught sample improved their understanding of electricity considerably more than the nontaught Ss, even the nontaught special class high scorers and gainers tended to score somewhat higher on the retest, indicating that they had learned from the prior testing and been able to apply it. (The possibility also exists that some members of the nontaught sample who were located at the school in which the classes were held might have been learning the materials informally from their taught friends). The nontaught nongainers however, did no differently, and in some instances, less well than on the pretest.

The remarkable aspect of the learning potential effect is the prevalence of its significance in the special-total regular class comparisons. The many positive learning potential F-ratios for posttest and gain absolute scores, pooling special and regular class children together, strongly suggest that learning potential status determined success in the unit for regular and special class students. The same situation was true when special and low achieving dull to normal IQ regular class children were compared on absolute scores. Learning potential effects accounted for significant proportions of the variance in consistency scores analyses. The fact that learning potential status was so unexpectedly strong as a determinant of success with regular and special class children indicates additional justification for seeing the overlap in abilities between these two populations.

Adjusting the posttest scores for the differences in IQs within the learning potential groups in each of the samples did not alter the pattern of results reported for the unadjusted scores. Learning potential status remained the major differentiating variable. With one exception, the linear component which predicts the level of ability displayed (high scorer, gainer, and nongainer, in that order) accounted for this effect, reaffirming that high scorers, gainers and nongainers, in that order, performed best on the electricity test following teaching. The lack of significant differences in scores between the taught total regular and special class samples indicated that the few differences in favor of the nonEMRs, which were apparent on the unadjusted scores were a function of differences in IQ between the students in the two samples.

In summary, several statements relevant to the intentions of this research can be made:

1. Certain special class children can perform equally well in educational settings as regular class children, when given material that is highly motivating and that compensates for the reading deficit. The hypothesis that high scorer and gainer special class children can be considered educationally rather than mentally retarded was borne out by the similarity of their learning pattern to that displayed by the dull and average IQ children who have done poorly in school. Though the special class children did know less about electricity to start with (or displayed more initial discomfort with, and avoided the evaluation task on the pretest), increased familiarity with the materials and the evaluation procedure, and exposure to the curriculum obscured these initial differences. The evidence from this study of classroom learning suggests that the rate of learning of the higher learning potential groups is at least equal to that displayed by the low achieving regular class peers, at least in these special types of non-verbal learning situations.

2. The analysis of the absolute scores attained by the special class nongainers indicated a very small increase in their understanding of electricity following teaching. Their consistency scores also indicated a negligible increase in their ability to apply electricity principles consistently across items within a set. The data from this study, then, lends additional support to the thesis that the nongainer in special class is handicapped in his ability to learn and profit from experience as well as in his reading prowess, a deficit he tends to share with his classmates. The differences in the performances characteristic of these children as opposed to high scorers and gainers suggests that nongainers are functioning as mentally retarded children, whereas the former groups appear to be more clearly educationally retarded.

3. The strong learning potential effects in all the data validate the learning potential task as a diagnostic instrument uncovering ability to learn and profit from experience, which is our working definition of intelligence in its broader implication. Since the Kohs block design task relies on first hand experience as a mode of learning and the manipulative exploratory Batteries and Bulbs unit does likewise, it is logical that high scorers outperformed gainers who in turn outperformed nongainers. The strong learning potential effect across special and regular class groups, evident when no class effect between groups exists, is most interesting, indicating as it does that successful performance in a nonverbal manipulative unit may be determined not by a verbally based IQ score, the score that initially segregates the special class student, but by learning potential status or performance on related nonverbal tasks. For special class children pre-, posttest and gain scores from the Electricity Evaluation Test

were correlated with nonverbal measures such as the Raven's Progressive Matrices and the WISC Performance Scale IQ but not with reading achievement and Stanford Binet scores.

The major cognitive strengths of both special and low achieving regular class students, pooled by learning potential status rather than class placement, would appear to be in the areas of learning that depend minimally on verbal conceptual and expressive skills, including reading. Successful educational experiences reflected in achievements at or above grade level probably results in better problem solving abilities more generally, but there is probably considerable overlap in problem solving abilities when the tasks do not require appreciable verbal - conceptual or expressive skills as is reflected in the correlational data and in performance on the electricity unit. For the child who is performing poorly in school, this evidence of his ability to learn and profit from experience (i.e., intelligence) may be critical in demonstrating that the child functions more satisfactorily than his school performance or scholastic aptitude score (Binet IQ) indicates. This "contrary" evidence must then be employed to re-engage these students in meaningful learning activities.

Several educational implications follow from the results:

1. Since certain EMRs show themselves to be very capable when taught with nonverbal exploratory-manipulative material it may be that more such units, in the field of science, mathematics, and/or social studies should be adopted for this type of learning. Further research is of course needed on whether the manipulation and consequent first hand experience with the materials was the crucial factor in expediting learning in the electricity unit, or whether the novelty of the subject matter, its relevance to the student's interest and the fact of the minimization of reading were more important. Such research has practical implications as well as a theoretical one, since supplying materials to each student for these types of units is an extremely expensive undertaking for a school system. As a concluding study for this project, a pilot study comparing the exploratory-manipulation and lecture-demonstration version of this same electricity course was carried out using only special class subjects. Results and further discussion of this problem are found in Section VI.

2. Just as high scorer and gainer students with IQs in the educable mentally retarded ranges may require novel presentations and interventions by which to learn that will tend to minimize the negative effects of their poor literacy skills, so, too, would low achieving regular class children with dull to average IQs seem to be able to profit from similar types of specially designed classroom interventions. If the implications of the learning potential hypothesis are correct, then skill in literacy, however small, and the consequent ability to profit, however minimally,

from the reading laden curricula favored in current schools may be the only factor which distinguishes the regular class school failure from the gainer and high scorer. Yet if the low achiever can learn considerably from non-verbal presentations as was demonstrated in the present study, he should be given the opportunity to work with these types of curricula so he can be meaningfully re-engaged in school learning. Again, research is needed with the low achiever to determine which type of curriculum he actually does profit best from.

3. Gainers and high scorers might do better academically in a regular class placement than in a special class. If (a) certain changes were made in curricula to compensate for their reading deficits and/or (b) they were given extensive work in reading to try and bring their skills up to regular class standards. It could be that if gainers and high scorer EMRs were not in special class, with its accompanying stigma of being "dumb", or unable, or slow to learn, they might be as capable of learning and applying new concepts as the regular class students. This regular class placement, with appropriate modifications that would individualize learning opportunities for these high risk students, might result in considerable alleviation in their school difficulties. These changes might require a higher proportion of nonverbal learning opportunities such as this laboratory science unit, and others like it in science, mathematics, and social studies. It would also require considerable expenditure of effort in remedial work in language arts subjects to determine whether these students can improve their competence markedly. For these severe school failures, it might also require explorations of the utility of providing tangible rewards or incentives to further motivate them to work in the language arts subjects in which they have experienced considerable failure. It may be that within the context of successful work in the nonverbal academic presentations and tangible rewards for trying to achieve more satisfactorily with the negatively loaded language arts materials, these students' motivation to work productively in school might be increased considerably, and result in markedly higher levels of achievement. However, a regular class placement for the non-reader based on a learning potential rating in which he would be left to fend for himself without special attention, would only do him a disservice, and continue the debilitating pattern of failure.

It may be, however, that the "true" reason for the pattern of severe school failure and the low scholastic aptitude score which resulted in the special class placement is a disability in efficiently processing verbal-conceptual materials. If there are suspicions of this type of disability, intensive efforts must be expended to determine whether it is an intrinsic deficiency or a function of poor prior experience at home and in school. If these intensive efforts do not yield an appreciable upgrading of skills, then this failure may point to the major deficit or defect which may be typical of the child who is classified as educable mentally retarded. Alternate programs of educating these able but nonverbal children would then have to be formulated which would direct themselves toward maximizing these children's strengths in the nonverbal areas, and seek to have

them obtain a maximal level of proficiency in the verbal-conceptual areas, such as reading. It should be self-evident that success in any of these types of determinations will be heightened by interviewing at younger chronological ages than those of the students participating in this study so that the negatively reinforcing cycle of failure might be short circuited before it becomes established for the child.

The study as projected in the proposal envisioned a comparison of the special class students' performance with mental age (MA) controls. But the need for this comparison was predicated on evidence of the inferior performance of the special class students to that of the CA controls. However, the very consistent findings that special and regular class students, especially the low achievers, performed no differently following teaching, and the critical importance of the learning potential status variable in determining score level suggested that the basic intent of the study had been fulfilled. This study basically sought to demonstrate differences in educability by learning potential status within a psychometrically homogeneous EMR group. It would then provide further validity for a learning potential assessment procedure, and substantiate further the educational implications of this procedure, namely that the high scorer and gainer special class child is educationally not mentally retarded. The consistent results meant that a comparison with MA controls would not add to the argument. The sole additional contribution of an MA contrast would be the finding that their performance level might be similar to that of the nongainer group, a pattern suggested by Budoff and Pagell's (1968) results with a concept sorting task. However, in light of the very small gain demonstrated by the nongainers following teaching, the same pattern of nonlearning displayed by the nongainers could scarcely be expected of MA regular class controls. Again, the nongainers' failure to improve their understanding of electricity appreciably following teaching suggests a functional inability which cannot be explained by slow development, or mental age, but probably by an intrinsic defect, a <u>mental</u> retardation.

What factors besides learning potential determined success in the curriculum and did they vary for special and regular class children? The data from the correlations of observation measures (made by both the teacher and by two independent observers) shows that the two groups did differ. Apparently, the special class child's evident involvement with the work bears less relationship with what he is actually learning than was true for the regular class student, at least to an outside observer. The teacher, who worked intimately with the students correctly gauged the special class children's progress in understanding regardless of their work habits, but the outside observer only caught their seemingly non-work oriented habits such as aimless manipulation of the materials, or stretches of boredom following periods of interest.

Need for acknowledgement, the other measure rated by outside observers, bore no relationship at all to the posttest and gain scores of either special or regular class. Cooperation, the last teacher variable, likewise was unrelated to either special or regular class posttest or improvement scores, suggesting that both for EMRs or dull to average IQ regular class children, "good" behavior bore little relationship to the actual learning process that was taking place within the child. A docile student who ranked high in "Cooperation" may be one who passively responds to the behavioral demands of the teacher while allowing the better part of his mind to wander completely away from the learning material at hand. On the other side, the provocative student with the obviously short attention span may actually be concentrating and taking in material in those few minutes he is being "cooperative".

When various school related and psychometric variables such as achievement tests, IQ scores, Raven's Progressive Matrices, etc., were correlated with selected electricity pretest, posttest and gain scores, the pattern of significant results indicated that only the nonverbal measures were related to success in the electricity unit. There were very few significant correlations between the school and psychometric measures and posttest and gain scores for regular class children. More significant correlations are evident on posttest and gain scores for special class students and, second and third administrations of the block designs and Raven's percentile rank are related to practically all of the selected variables and arithmetic achievement score.

What is interesting is that electricity scores following teaching and the nonverbal measures were correlated. Thus the high learning potential EMRs were best described in terms of their performances on these other measures by their attainments following a period in which they became familiar with, and learned about electricity. Also interesting was that none of the significant correlations with the electricity scores were with the verbally based measures such as reading achievement or the Stanford Binet IQ. In general, these findings and the fact that significant correlations did appear between the electricity scores and nonverbal measures supports the premises of the argument for a learning potential procedure. The student EMR who is likely to profit from a nonverbal manipulative curriculum like the Batteries and Bulbs electricity unit is the one who exhibited strong reasoning ability on nonverbal tasks, not on tests like the Stanford Binet which are heavily dependent on verbal-expressive or conceptual skills.

The underlying ability of the high learning potential EMR became evident following teaching. They were not evident in the pretest scores. Motivationally, too, posttest scores correlated most highly with a variety of positive factors. The more successful special class student on the posttest expressed high value for school, and as not being completely defeated by his lack of progress in school. He even continued to see school as being relevant to his future. Higher scores for social and academic responsibility were also correlated with posttest, but <u>not</u> pretest scores. Thus, among

special class students, certain statements derived from the interview related to school performance were best validated by high correlations with posttest scores, i.e., following an opportunity in which the EMR can obtain a sense of competence with the subject matter of electricity, rather than with initial level of performance with these materials.

These findings suggest additional support for a learning potential type of assessment procedure. It may be that a prior successful history of successful learning in school means that present level of functioning in a more or less novel situation, such as the electricity pretest, relates summarily to responses in other life areas for the regular class student. But the stigmatized special class child's hopes, which are presumably embodied in their interview responses, seem to be more accurately related to "what they might do if . . ." This is the situation represented by the posttest scores in this study. If these different relationships for special and regular students can be validated with other motivational variables, these findings would constitute additional support for the learning potential type of assessment procedure with severely school failing youngsters whose low Binet IQs place them in the educable mentally retarded range.

VI. A COMPARISON OF SPECIAL CLASS STUDENTS' PERFORMANCE ON TWO PARALLEL ELECTRICITY UNITS: EXPLORATORY-MANIPULATIVE VS. LECTURE-DEMONSTRATION.

A. Introduction.

The major study clearly demonstrated that the performance of high learning potential educable special class students was similar to that of regular class students, especially low achievers, when both groups were involved in a laboratory science electricity curriculum. The unit in the major study was exploratory-manipulative; it had been designed to utilize the much advertised concrete learning needs of slow learners as well as to provide a rough parallel to the manipulation-becoming-internalization learning that is hypothesized to take place on the Kohs task. The similarity in performance levels of the special and regular class students following this unit was attributed to the de-emphasis on reading and verbal expression, and maximization of the opportunities to learn by handling and exploring for solutions with the materials themselves, with a consequent increased interest and motivation. However, this argument would be stronger if there had been an additional condition in which the students could have learned the same electricity "facts" without the exploratory-manipulative experience, but with a unit that was presented in a lecture demonstration format.

A small scale comparison of these two ways of teaching electricity to adolescent EMRs was therefore planned as a concluding study for this project to serve as preliminary test for future work. The adaptation of the Batteries and Bulbs unit used in the major study with its heavy emphasis on pupil-oriented experiential learning would be contrasted with a lecture-demonstration course on the same material. The latter course, directed by the teacher, would of necessity be more structured; a logical ordering of concepts, illustrated by examples from Batteries and Bulbs, would replace the direct experiencing of electrical facts as the means by which the students could organize their acquired knowledge. The course would differ from a traditional science class in that it would still minimize formal electricity theory and its terminology, and reading and written exercises. It would further differ from an ordinary science experience in that the minimal laboratory experience usually permitted in the conventional science classroom -- occasional practice experiments and assisting the teacher in demonstrations -- would be completely eliminated. In order to control for the manipulation variable in the comparison of the two curricula, children in the lecture-demonstration group would be permitted at no time to handle any equipment.

Following from the results of the major study, it was expected that the manipulation group would learn significantly more about electricity than the lecture-demonstration group because the opportunity to learn through direct experience with the materials would be better suited to their natural learning style. It was also thought that learning potential would be a significant determinant of success in the manipulation group but would have no similar effect in the lecture-demonstration group. This hypothesis was derived from the fact that the main determinant of learning potential, the Kohs block design assessment procedure, is a manipulative nonverbal task, and that the manipulation unit would offer an opportunity for the same kind of learning as the Kohs. The

lecture-demonstration unit, on the other hand, would not depend on this particular style of learning; it would offer opportunity instead to employ skills of listening, observation, and ability to discuss phenomena. One would therefore expect learning potential to be a good predictor of final performance level in the manipulative group and not in the lecture-demonstration class.

B. Methods and Procedure.

1. Description of the manipulation and the lecture-demonstration units. In the autumn of 1967 the dimensions of the two comparison units were fully defined and the lecture-demonstration unit was written.* The manipulation unit was essentially the same core unit previously described in section III. By experimenting with their own equipment, children were to make and learn about complete circuits, parallel and series bulbs and batteries circuits, resistance, etc. The teacher would structure the classes by setting out problems and project areas for the lessons, providing directions for the children to proceed in, and materials with which to work. Outside of these strictures, however, the class would be pupil rather than teacher centered. Emphasis would be on the individual work the students would produce, and learning would take place through their own exploration with the materials. Visual aids such as picture notebooks which would offer ideas and provide a place for students to record results, and flashcards for review and occasional drill would be employed, but their use would be supplementary to the actual experimentation.

The lecture-demonstration unit as planned would start out with several introductory lessons on the importance of electricity, the difference between static and current electricity and "how electricity travels", magnetism and the relationship between magnetism and electricity. After these introductory lessons, conducted through teacher lecture and demonstration and teacher-pupil discussion, the teacher would take apart a battery. He or she would demonstrate how and when the battery could or could not work, and teacher and class would then discuss complete circuits. In this last discussion, the teacher would demonstrate the different kinds of simple complete circuits and drill the children in identifying them. After this, the topics of resistance, conductors, and insulators would be presented -- frequent demonstrations of how different materials affected the flow of electricity and hence the brightness of the bulb would be made. Later, parallel and series batteries and bulbs first in simple and then complex circuits would be studied in detail, again employing the teacher lecture, demonstration, and class discussion method. (See Appendix H, which presents the teacher's work book).

*The major portion of this task was completed by Mrs. Karen Corbett Howell, with the consultation of the second author.

In the obviously teacher-centered classroom of this second method, learning would depend on listening carefully and above all, on watching the demonstrations. To secure pupils' attention to the subject matter itself without allowing them to handle the materials, more emphasis would have to be put on visual aids than in the manipulation group. A picture chart to illustrate all points the teacher would make would be provided. Students would have notebooks with many replications of the chart pictures as well as many picture exercises to work out ("Trace the flow of electricity in these set-ups"). Numerous flashcards would be provided for each lesson for teaching as well as drill work, and finally, three dimensional boards of active batteries, bulbs, and wires would be used to demonstrate new ideas, and to review points from previous lessons. (Appendix I presents the student's work book.)

To summarize, the features that would distinguish the lecture-demonstration unit from the manipulation unit were 1) absolute lack of manipulation of materials on the part of the students, 2) more highly structured classes with the teacher leading all lectures and discussions and doing all demonstrations, and 3) greater emphasis on class discussion and drill exercises in notebooks and during discussions.

Efforts were made to minimize any lesser differences between the units that might complicate interpretation of the results. Time exposure would be kept equivalent. It was decided that the lecture-demonstration unit could not sustain children's attention longer than 18 class periods and the curriculum was thus written to be presented three times a week for six consecutive weeks. The manipulation unit was pruned down from its original 26 lessons to 18 in order to match the length of the lecture-demonstration intervention. Besides being an excellent visual aid, the demonstration boards with their three dimensional battery and bulb set-ups would be used to provide equivalent visual exposure to the elements composing the evaluation of the unit in both classes and thus insure the fairness of testing with real set-ups in both groups.

A final addition to both units was the printing of the legend "Grade 8" on the title page of the two curricula notebooks. In the main study some students were suspicious that the special unit they were receiving was "baby material" that would further stigmatize them as slow learners. The words "Grade 8" were added to the notebooks in the concluding study to increase the students' confidence in the curricula by assuring them that they were receiving bona fide junior high school material.

2. Design of the experiment.

The experiment was similar in design to the major experiment. Pretest scores were obtained on a special evaluation measure, the curriculum interventions administered and posttest measures obtained. The difference was that two curriculum interventions were to be used instead of one. Two experimental classes were involved -- one allocated to each teaching condition. Nontaught controls were not included in the design since those included in the major study were still available as a comparison group. The subject sample was limited to special class EMRs, due to limitations of time and budget.

3. Evaluation.

The same evaluation instrument was used in this study as that in the major study, described in section IIC. Again, the test had two major divisions. The first tested knowledge about electricity and the ability to apply it, through having the student answer questions about real three dimensional set-ups mounted on masonite pegboards. One subsection of the Boards section of the test dealt with simple circuits, and two subsections dealt with a combination of both simple and complex circuits. The second division tested the same factors as the first plus the ability to translate to a more abstract mode; S answered similar questions about set-ups as in the Boards section of the test, but the setups were presented this time in both realistic and schematic diagram form. There were two subsections of realistic diagrams, one dealing with simple and one with both simple and complex circuits, and one subsection of diagrams in this second division of the test.

The format in both divisions, boards and diagrams, was a simple multiple choice test in which the S had to choose the correct answer from among four alternatives. The basic questions for subsections dealing with simple circuits alone in both the Boards and Diagrams divisions was "Which bulb, among these four circuits, will light?"; that in the subsections dealing with a mixture of simple and complex circuits in both Boards and Diagrams was "If this bulb would light, would it be the same as standard brightness[2], brighter than standard, dimmer or wouldn't it light at all?" or a variation of this. The entire test was individually administered (See section IIC, for a more complete description of the test).

One major section was added to the evaluation scheme for this pilot study. Its purpose was to see if the students knew the reasons why electricity functioned as it did in different set-ups; it acted as a replacement for the "why" questions asked about particular set-ups in the major (1967) study. In this section, also administered individually, S was presented with pictures of a set-up with four statements following each one, and asked to choose the statement which best fit the picture. For example, on an item showing two series batteries and one bulb, S was asked to "choose the statement which best fits the picture" from among the following four: 1. The wire between the batteries short circuits the power. 2. Each battery gives out only one-half of its electric power. 3. The two batteries cancel each other out, and 4. The two batteries add their power together. E read the choices to S while S listened, followed each sentence while it was read, and then made his choice. There were 33 items in this section with four appropriate choices for each picture. (See Appendix F)

[2]Standard brightness, it will be recalled, was defined as the brightness of a bulb in a circuit composed of one bulb and one battery.

In summary there were eleven scores to be derived from the revised electricity test. As in the evaluation instrument used in the major study, in the Boards section: 1. Simple Circuits, 2. Complex Circuits, 3. Complex Circuits II, 4. a subtotal for the Boards section; in the Diagrams Section, 5. Simple Circuits, 6. Complex Circuits, 7. Circuits represented in Schematic Diagramming, or "Schematics", 8. A subtotal for the Diagrams division, 9. A subtotal of the two subtotals of the Boards and Diagrams divisions. The additional scores were: 10 The Reasons section, 11. A grand total (#9 plus #10).

4. Hypotheses.

Stated in terms of the sections of the test, the formal hypotheses of the experiment were:

1) Measured by gain from pretest to posttest, the manipulation group would have higher improvement scores than the lecture-demonstration group on all sections of the test except section (10), the reasons section. Posttest scores of the manipulation group would also be higher than the lecture-demonstration group.

2) The posttest and improvement scores from pre- to posttest on section (10), the reasons section, would be higher in the lecture-demonstration group because this group would have greater exposure to verbal explanations of electricity in their class and hence be able to associate reasons with set-ups better than the manipulation group.

3) There would be a significant effect of learning potential operating in the posttest and improvement scores of the manipulation group for all sections except the verbal reasons section (10). Since the style of learning in the manipulation class would parallel that required in the basic learning potential task, success in one should therefore imply success in the other. In the lecture-demonstration group such an effect would not operate.

4) When scores of the 1968 manipulation group were compared with the 1967 manipulation group (EMR) there would be no significant differences in mean pretest, posttest, or gain scores. Both groups would be higher than the 1968 lecture-demonstration group in posttest scores and improvement scores.

5) All groups, 1967 manipulation, 1968 manipulation and 1968 lecture-demonstration would have significantly higher posttest and gain scores than the 1967 nontaught controls.

5. Description of sample.

The Ss used in this study were 26 special class students from the two special classes attending a junior high school in a predominantly Italian-American section of Boston. Originally all 32 students in these classes were assigned to either the manipulation or the lecture-demonstration condition, but six were eventually dropped because of excessive absence or the shift of some Ss to another school.

The 32 students were administered the Khos learning potential procedure (see Section III) and designated as high scorers, gainers, and nongainers. Sixteen pairs of students were then matched on the basis of learning potential status, chronological age and latest IQ available at that time. The loss of six students eventually caused a slight imbalance in the groups. The final breakdown by learning potential status and sex for each class was : manipulation group: high scorer-3 (3 male), gainer-5 (2 male,3 female), and nongainers-4 (2 male, 2 female); and lecture-demonstration group: high scorer-5 (4 male, 1 female), gainer-5 (3 male, 2 female), and nongainer-4 (3 male, 1 female). During the course of the intervention new Wechsler Intelligence Scale for Children (WISC) IQ scores were obtained on 15 of the 26 Ss that had been tested more than two years prior to this experiment. The new scores were sufficiently different that the earlier match for IQ was disrupted and an IQ imbalance in the samples resulted.

Table 29 presents the means and standard deviations for the two final treatment groups subdivided by learning potential status, for latest Binet or WISC Full Scale IQ (administered after the intervention began), latest verbal IQ (WISC verbal or Stanford-Binet score), chronological age at the start of the experiment, and occupation of the principal wage earner of the family (a measure of social class).

Two way analyses of variance (treatment, 2 levels; learning potential status, 3 levels) were performed on each of these measures to determine whether there were any differences among the treatment and/or learning potential groups on the different parts of the test. Table "29" also presents a summary of the F-ratios for the treatment, learning potential and the interaction effects for each of the control variables. The treatment groups were similar except on latest verbal scale (VIQ) or Binet IQs. However, since the higher VIQs were present in the manipulation rather than the lecture-demonstration groups, the bias was considered to be less pernicious to the hypotheses of the study than if the higher VIQ students had received the more verbal unit.

Differences on both IQ measures due to learning potential status were even more evident. Recent studies of large numbers of special class students compared by learning potential status indicated this same difference, namely, that high scorers have higher IQs than gainers and nongainers, and gainers' IQs tend to be higher than those of nongainers (Budoff, 1968). The requirements of scheduling, and the few students available within the school allotted to the study, meant that some IQ imbalance could not be prevented. What was more important for purposes of this study than these learning potential differences, however, is that there were no significant interactions between treatment group and learning potential status on any of the four measures. Thus, we may conclude that the learning potential groups were equivalent across treatment groups in chronological age, occupation of the principal wage earner (social class), IQ and VIQ,i.e, high scorers in both treatment groups have essentially similar IQ's, etc.

Table 29

Means, SDs and F-Ratios for Manipulative and Lecture-Demonstration Samples for IQ, VIQ, CA, and Occupational Rating of Principal Wage Earner

	df	Binet or WISC Full Scale IQ	Binet or WISC Verbal Scale IQ	CA	Mean Occupational Rating (Social Class Measure)
MANIPULATIVE					
High Scorer (N=3)					
Means		82.67	84.67	169.67	1.50
S.D.		4.04	8.08	8.62	1.50
Gainer (N=5)					
Means		72.60	73.80	179.60	2.40
S.D.		9.40	2.95	7.89	0.55
Nongainer (N=4)					
Means		68.25	68.75	168.75	2.25
S.D.		4.86	3.86	15.11	0.96
LECTURE-DEMONSTRATION					
High Scorer (N=5)					
Means		79.20	76.60	174.40	1.50
S.D.		4.15	4.72	13.90	1.50
Gainer (N=5)					
Means		76.20	72.20	174.20	2.20
S.D.		5.22	7.16	6.72	0.89
Nongainer (N=4)					
Means		61.75	63.75	171.75	1.50
S.D.		3.40	4.99	15.11	0.58
F-RATIOS	df				
Treatments	1	---	4.60[1]	---	---
Learning Potential Status	2	14.62[2]	12.40[2]	---	1.35
Tr x LP	2	1.83	---	---	---
MEAN SQUARE					
Subjects within Groups	20	33.45	28.95	136.07	1.06

[1] $p < .05$

[2] $p < .01$

6. Procedures.

The revised electricity evaluation test was administered to the original 32 students prior to the start of classes before Christmas vacation, 1967. On January 4, the two classes began. Classes were scheduled to be held three times a week for six consecutive weeks. However, snow days and the difficulty of covering the material in eighteen lessons caused a prolongation of the schedule, and in all, 22 lessons were actually held over a nine week period (a one week vacation period was included).

A major uncontrolled variable in the study was that two different teachers taught the two classes. Only two classes were available for inclusion in this study -- one for each condition. The same teacher who taught all groups in the major experiment previously described was scheduled to teach the manipulation group, but it was felt that she would not be comfortable teaching a group in the different style called for in the lecture-demonstration unit. Another teacher, therefore, was employed for the lecture-demonstration class. Thus, in all the analyses and conclusions drawn, it is critical to remember that teachers and treatment condition were confounded.

The pattern of the manipulation class was much the same as in the major study of 1967. The same teacher, Mrs. Jean Rosner, conducted the group, though the teaching assistant was not used this time in order to reduce unnecessary confusion in the class. The activities of the 1968 manipulation class, however, tended to be more varied and individualized than those of the 1967 one, and it is thus harder to estimate each child's exposure to the different topics. A rough summary of the time the class as whole concentrated on the separate areas follows: Complete circuits, Incomplete circuits and Short circuits: seven lessons; Parallel and Series Bulbs and Batteries Circuits: eight lessons; Resistance, with various wires: two lessons; Resistance, with two types of bulbs: two lessons; Conductors and Insulators (also a Resistance topic): two lessons; Review: 1 lesson.

The lecture-demonstration group, conducted by Mr. Robert King, basically followed the schedule prescribed by the written unit (see Appendix H). Five lessons were spent on background material including magnets and magnetism and how current electricity works before the actual introduction of chemical electricity, batteries, bulbs and wires. Two lessons were then spent on taking apart a battery and on complete circuits, and two more lessons explained resistance and showed applications of resistance in conductors and insulators and different wires. A lesson on short circuits and a review followed this. The remainder of the 22 lessons, or about half of the class time, was spent on parallel and series batteries and bulb circuits, presented in increasingly more complicated circuits. Throughout the whole unit, students did a great deal of exercise drill using flashcards and diagrams in notebooks. The unit concluded with a review, employing these aids as well as the demonstration boards.

One important addition to the unit, suggested by Mr. King, was a point-reward system for good questions and answers in class discussion and good work in oral and "written" exercises. This system started out as a fairly traditional teacher aid but was used so extensively in the lecture-demonstration group and became so important to the children, that its use may very well have affected the results of the study. A similar system for good experiments as well as good questions and answers was set up for the manipulation group, but it did not become nearly as prominent a feature of this class as in the lecture-demonstration group.

During the intervention, both teachers rated each of their students weekly on different behavioral measures. The teacher of the manipulation group rated the students on Ability, Productivity and Application (or work accomplished), and Cooperation, while the teacher of the lecture-demonstration group rated his pupils on Ability, Cooperation, Interest and Participation in class. Both rating scales ranged from 5 (low) to 1 (high). (Copies of these scales are found in Appendix D). It was supposed that there would be a high correlation between these measures and performance on posttest and improvement scores in both classes. No observation by independent observers was conducted.

The last classes were held on March 4, 1968. All posttests were administered in the week following the termination of classes.

C. Results

1. Analysis of pretest scores.

Analyses of variance were performed on the eleven pretest scores treatment (2 groups), and learning potential status (3) to determine the equivalence of the groups' initial knowledge of electricity. Table 30 presents the means and standard deviations from each of the groups on the eleven different sections of the test and summarizes the F ratios. Examination of the F-ratios shows that only one out of a possible thirty-three is significant -- the interaction effect of treatment and learning potential on Complex Circuits II in the Boards division of the test ($p < .05$). To obtain a finer measure, an analysis of covariance was run on the same pretest scores using three variables that might be expected to affect the scores and were not entirely controlled for in the design -- number of absences of each student, sex, and latest IQ. The results of this analysis when the three covariates were taken together again showed only the interaction effect on Complex Circuits II in the boards division of the test to be significant ($p < .01$). These two analyses of the pretest scores indicated the equivalence of initial level of knowledge in the two treatment groups, the three learning potential groups and the learning potential groups within the treatment groups.

2. Analysis of posttest and improvement scores.

Table 31 shows the posttest means and standard deviations for the eleven scores of the test and the summary analyses of variance performed

Table 30: Means, SDs, and F Ratios (no Covariates) on Electricity Pretest Scores for Manipulative and Lecture-Demonstration Samples

	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Boards & Diagrams Total	Reasons	Grand Total
MANIPULATIVE												
High Scorer (N=3)												
Means		2.00	10.67	2.67	15.33	4.00	10.33	8.33	22.67	38.00	11.00	49.00
S.D.		1.00	1.53	1.53	1.53	1.00	0.58	2.52	3.06	2.65	2.00	4.36
Gainer (N=5)												
Means		2.60	9.40	4.00	16.00	3.00	8.80	6.80	18.60	34.60	8.20	42.80
S.D.		1.34	2.30	1.58	3.32	1.73	1.79	1.48	1.95	3.85	3.42	4.44
Nongainer (N=4)												
Means		4.00	10.25	4.50	18.75	3.00	10.75	5.25	19.00	37.75	8.25	46.00
S.D.		1.41	3.30	1.00	4.86	1.41	4.43	1.71	5.35	9.36	0.50	9.27
LECTURE-DEMONSTRATION												
High Scorer (N=5)												
Means		2.60	11.20	4.40	18.20	2.60	9.20	6.20	18.00	36.20	9.00	45.20
S.D.		0.55	3.11	.89	2.95	1.34	2.49	1.92	4.95	5.68	1.87	6.98
Gainer (N=5)												
Means		2.60	9.80	4.00	16.40	2.40	10.20	6.40	19.00	35.40	9.00	44.40
S.D.		1.34	3.77	1.58	4.39	1.67	3.77	1.67	4.36	7.69	2.24	6.69
Nongainer (N=4)												
Means		3.50	10.50	2.25	16.25	3.00	5.75	6.75	15.50	31.75	9.00	40.75
S.D.		1.29	2.08	0.96	3.40	0.82	2.99	0.50	3.70	6.13	1.41	5.50
F RATIOS												
Treatments	1	--	--	--	--	1.32	1.29	--	2.04	--	--	--
Learning Potential Status	2	3.10	--	--	--	--	--	--	--	--	--	--
Tr. x LP	2	--	--	4.61^{1}	1.05	--	2.53	2.25	--	--	1.07	--
MEAN SQUARE												
Subjects within Groups	20	1.43	8.36	1.68	13.31	2.02	9.03	2.85	16.74	41.34	4.78	41.98

$^{1}p < .05$

$^{2}p < .01$

Table 31: Means, SDs, and F Ratios (no Covariates) on Electricity Posttest Scores for Manipulative and Lecture-Demonstration Samples

	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Boards & Diagrams Total	Reasons	Grand Total
ANIPULATIVE												
High Scorer (N=3)												
Means		5.00	18.00	5.00	28.00	9.00	16.67	12.00	37.67	65.67	15.67	81.33
S.D.		1.00	1.73	1.00	3.00	1.00	3.22	1.73	4.04	2.08	3.06	2.52
Gainer (N=5)												
Means		4.60	12.60	5.20	22.40	5.40	13.20	8.80	27.40	49.80	9.60	59.40
S.D.		1.82	3.29	1.64	5.86	2.30	1.64	3.42	7.06	11.56	1.34	12.50
Nongainer (N=4)												
Means		4.50	11.75	5.00	21.25	4.25	11.75	8.00	24.00	45.25	8.00	53.25
S.D.		.58	6.70	1.41	7.72	1.71	1.71	2.58	5.10	12.66	3.16	15.46
ECTURE-DEMONSTRATION												
High Scorer (N=5)												
Means		5.80	15.40	6.80	28.00	7.40	14.00	11.40	32.80	60.80	11.60	72.40
S.D.		3.96	5.13	1.92	9.30	2.97	2.00	1.14	3.70	12.15	2.88	13.94
Gainer (N=5)												
Means		5.80	13.00	4.80	23.60	6.20	12.40	9.80	28.40	52.00	10.80	62.80
S.D.		3.19	3.32	1.10	5.37	4.38	3.21	2.95	9.40	14.65	4.32	18.74
Nongainer (N=4)												
Means		2.25	9.00	4.00	15.25	3.25	6.00	5.50	14.75	30.00	9.00	39.00
S.D.		1.50	2.16	1.41	3.40	1.26	1.16	0.58	0.96	4.24	1.83	5.48
RATIOS												
Treatments	1	--	--	--	--	--	10.59[2]	--	2.75	1.45	--	1.17
Learning Potential Status	2	1.71	4.21[1]	2.51	4.58[1]	4.86[1]	15.25[2]	8.36[2]	12.98[2]	9.91[2]	5.05[1]	9.68[2]
r x LP	2	1.18	--	1.90	--	--	2.70	1.20	1.65	1.34	2.04	1.03
EAN SQUARE												
Subjects within Groups	20	6.33	17.36	2.22	41.50	7.44	5.07	5.70	36.03	126.35	8.69	181.33

[1] $p < .05$
[2] $p < .01$

independently on each of these scores, for the treatment, learning potential and interaction between the treatment and learning potential factors. Table 32 gives the same data and results for the improvement scores (posttest minus pretest scores).

Teaching Treatments - Exploratory-Manipulative vs. Lecture- Demonstration. When the effects of treatment are considered, the notable result is the total absence of significant F-ratios among the improvement scores and the presence of only one significant ratio among the posttest scores (Diagrams-Complex Circuits). Thus, there is no reason to reject the null hypothesis that the improvement scores of the two treatment groups were equivalent. These results fail to support the major hypothesis of the study: that the manipulation students will perform significantly better than the lecture-demonstration group on the nine posttest and improvement scores.

The second hypothesis predicted that the lecture-demonstration group would perform significantly better than the manipulation group on the verbal reasons score. It was also not supported by the results. The posttest and improvement scores for the reasons' subscore show a low level of performances by all treatment and learning potential groups, and the treatment F-ratios were not significant.

Learning Potential Factor. The learning potential effect in the analyses of variance of the improvement (and posttest) scores showed far different results. Eight of the eleven F-ratios for the improvement scores were significant -- except for Boards-Complex Circuits I, Boards Complex Circuits II and the verbal reasons subscore. Among the eleven F ratios for the posttest only Boards-Simple Circuits and Boards-complex Circuits II were nonsignificant. Thus, a strong learning potential effect cutting across treatment groups was demonstrated. The means for the subscores of the posttest and improvement scores fall in the expected direction -- high scorers in both treatment groups performed better than gainers, who in turn performed better than non-gainers. Further, no learning potential or treatment interactions were significant.

The results would seem to indicate that learning potential status is a factor which is generally applicable to functioning on both types of curricular presentations. The third hypothesis sought to predict that learning potential status would be most influential in determining performance level on the manipulative unit because of the obvious parallels with the Kohs Block Design assessment task. But it was not supported, which was both surprising and intriguing, suggesting that the ability to reason and learn inductively, indicated by high learning potential status, is a more pervasive characteristic than was previously suspected.

3. Analyses of Covariance.

The lack of control for latest IQ and the lack of control for sex and absences between the treatment groups suggested that the posttest and gain scores should be corrected by latest IQ, sex, and absence. Analyses of covariance were therefore run correcting the posttest and gain

Table 32: Means, SDs, and F Ratios (no Covariates) on Electricity Gain Scores for Manipulative and Lecture-Demonstration Samples

	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Boards & Diagrams Total	Reasons	Grand Total
NIPULATIVE												
High Scorer (N=3)												
Means		53.00	57.33	52.33	62.67	55.00	56.33	53.67	65.00	77.67	54.67	82.33
S.D.		1.73	.58	1.53	2.31	1.73	3.79	1.16	2.65	4.04	2.31	2.08
Gainer (N=5)												
Means		52.00	53.20	51.20	56.40	52.40	54.40	52.00	58.80	65.20	51.40	66.60
S.D.		1.00	3.70	1.10	3.36	2.51	2.55	2.55	6.98	9.47	3.44	9.34
Nongainer (N=4)												
Means		50.50	51.50	50.50	52.50	51.25	51.00	52.75	55.00	57.50	49.75	57.25
S.D.		1.73	8.06	1.73	10.47	1.26	3.83	4.11	5.16	14.89	3.10	17.91
CTURE-DEMONSTRATION												
High Scorer (N=5)												
Means		53.20	54.20	52.40	59.80	54.80	54.80	55.20	64.80	74.60	52.60	77.20
S.D.		3.90	6.91	2.51	9.91	3.11	4.09	2.78	7.33	15.55	4.04	19.28
Gainer (N=5)												
Means		53.20	53.20	50.80	57.20	53.80	52.20	53.40	59.40	66.60	51.80	68.40
S.D.		2.39	5.45	1.10	7.79	3.03	2.77	3.29	8.33	15.57	5.02	19.11
ngainer (N=4)												
Means		48.75	48.50	51.75	49.00	50.25	50.25	48.75	49.25	48.25	50.00	48.25
S.D.		0.50	4.20	1.71	4.97	0.96	3.69	0.96	3.10	5.56	2.00	6.99
RATIOS	df											
Treatments	1	--	--	--	--	--	1.30	--	--	--	--	--
Learning Potential Status	2	5.63[2]	1.91	1.70	3.80[1]	5.11[1]	3.84[1]	3.85[1]	8.06[2]	6.83[2]	1.85	6.34[2]
. x LP	2	--	--	--	--	--	--	2.62	--	--	--	--
AN SQUARE												
Subjects within Groups	20	5.17	30.65	2.86	54.72	6.72	11.81	7.81	40.48	154.28	13.23	220.77

[1] $<.05$

[2] $<.01$

scores of the different treatment and learning potential groups by these factors. When all three covariates were considered, in combination, no significant F-ratios were evident on either the posttest or improvement scores. In order to identify which of the three covariates had cancelled the learning potential effect evident in the original analysis of variance (if indeed it was one of the covariates by itself), three separate groups of analyses of covariance were carried out, each applying one covariate correction at a time. The analyses performed with the absence covariate and those performed with the sex covariate yielded F-ratios practically identical in pattern to those of the original analyses of variance, e.g., no treatment or interaction effects, but a strong learning potential effect. The analyses of covariance employing latest IQ as the covariate, however, displayed a pattern similar to that obtained when three covariates were applied in combination, i.e., there were no significant F-ratios on any of the factors.

It was therefore assumed that correcting the posttest and gain scores by the "latest IQ" covariate eliminated the effect of learning potential status. The "latest IQ" scores used for the IQ covariate consisted of both full scale WISC and Binet IQs. Fifteen students had been retested on the WISC during their period of participation in the curriculum unit. Data on large samples of special class students indicates that there are no differences between Stanford Binet and WISC VIQs by learning potential status. These scores invariably fall in the retarded IQ ranges. But there are appreciable differences between these scores and WISC Performance Scale IQs (PIQs) by learning potential status. That is, the mean IQs of high scorers and gainers was 88.80 and 80.69 respectively, while that of the nongainers was 68.86 (Budoff, 1968). Thus, whatever ability is being tapped by the learning potential assessment procedure, is also reflected in the Wechsler Performance Scale IQ. The VIQ score, like the Binet IQ, can be considered a measure of scholastic aptitude, and predicts these students' poor performance in school accurately. The higher WISC Full Scale IQs reflected the weighted contribution of the higher PIQs. Hence, it was hypothesized that the verbally based scholastic aptitude scores would not cancel out the learning potential effect but the Full Scale IQs might, since the latter are weighted for the higher PIQs.

The posttest and gain scores were reanalyzed, covarying the scores with a verbally based IQ measure (either the WISC verbal scale or the Stanford-Binet IQ). The results showed no significant ratios for the treatment effect, and only a few trends for the interaction effect on both sets of scores, ($p<.10$ or $<.20$). However, six of the eleven posttest scores and seven of the eleven improvement scores showed a trend for the learning potential effect ($p<.10$ or $<.20$). Since most of the F-ratios for learning potential in the analyses of variance without covariates were significant at the .05 or .01 level, differences in verbal IQ partially accounts for the differences among the learning potential groups on posttest and gain scores. The F-ratios for the IQ and VIQ covariate analyses are found in Table 33.

Table 33: F Ratios on Electricity Posttest and Gain Scores Covaried for IQ for Manipulative and Lecture-Demonstration Samples.

		Simple Circuits	Complex Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Boards & Diagrams Total	Reasons	Grand Total
POST TEST												
F Ratios Treatment	df=1	---	---	---	---	---	9.45	---	2.09	---	---	---
L.P.	df=2	---	---	---	---	1.39	3.07	---	2.37	1.38	---	1.37
T. x L.P.	df=2	---	---	2.32	---	---	1.43	---	---	---	1.94	---
MEAN SQUARE Subjects within Groups	df=20	6.32	16.90	2.06	38.47	7.75	4.77	4.90	33.66	114.09	8.32	[illegible]63.23
GAIN												
Treatment	df=1	---	---	---	---	---	---	---	---	---	---	---
L.P.	df=2	1.31	---	---	---	1.78	---	---	---	---	---	---
T. x L.P.	df=2	---	---	---	---	---	---	1.77	---	---	---	---
MEAN SQUARE Subjects within Groups	df=20	5.20	29.79	2.97	52.53	7.06	11.33	6.70	37.01	141.08	13.73	[illegible]06.74

It is interesting that the learning potential effect tends to disappear when the scores are adjusted with an IQ covariate, most particularly when the IQ scores used included performance types of tasks, such as Binet or VIQ of the WISC. Hence, the learning potential task does seem to tap abilities which are different from those which appear on the verbally biased IQ tests.

4. Correlations of electricity evaluation scores with observed classroom behaviors.

In order to determine the strength of the relationship between students' behavior in class and their performance as measured by the electricity evaluation scheme, correlations were run between the several ratings the teachers had made of the students' behavior in each class during the interventions, and pretest, posttest and improvement scores. The variables correlated with the scores were Ability, Productivity and Application to Work (or work accomplished) and Cooperation in the manipulation class, and Ability, Cooperation, Participation in Class and Interest in the lecture-demonstration class. Since the pattern of each child's performance on the separate variables tended to remain stable, the scores used for the correlations were the averages of the behavioral measures taken over the nine weeks.

Table 34 presents the correlations of these averages of teacher ratings with the posttest and improvement scores. As one might expect, there was only one significant correlation with pretest scores between both groups -- Ability with Diagrams -- Complex Circuits ($\underline{r}$ -.546, $p<.05$) in the lecture-demonstration group. Hence, these correlations were not included in the table.

The correlations with posttest and improvement scores will be discussed separately for the two variables recorded for each class -- (Ability and Cooperation), and for those variables measured in the manipulation or lecture-demonstration class alone (Productivity, Participation in class, and Interest). In both teacher ratings of classes Ability correlated significantly with several of the posttest and improvement measures indicating that in these particular special class groups, children's grasp of material was evident from their class performance in both a manipulation and lecture-demonstration situation. This was not the case with the Cooperation variable however. The low incidence of significant correlations with Cooperation seems to show that in neither type of class does an EMR adolescent's apparent behavior bear a strong relationship to the learning that is taking place within him. These results are similar to the findings in the major study.

One can make no generalizations across classes for the Productivity and Application variable, measured only in the manipulation group, and the Participation in Class and Interest variables, measured only in the lecture-demonstration class. The significant correlations with the Productivity variable are hard to interpret even within the manipulation group since few appear for the improvement scores and only one for the

Table 34. Correlational Matrix for Classroom Behaviors as Related to Electricity Posttest and Gain Scores for Manipulation and Lecture-Demonstration Students. (Significant r's only)

	Simple Circuits	Complex Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Boards & Diagrams Total	Reasons	Grand Total
POSTTEST											
Manipulation (N = 12)											
Ability	---	---	---	---	---	-.582[1]	---	-.615[1]	-.610[1]	---	-.607[1]
Productivity & Application to Work	---	---	---	---	-.607[1]	---	---	---	---	---	---
Cooperation	---	---	---	---	---	---	---	---	---	---	---
Lecture-Demonstration (N = 14)											
Ability	-.673[2]	---	---	---	-.662[2]	-.662[2]	-.825[2]	-.790[2]	-.700[2]	---	-.682[2]
Cooperation	---	---	---	---	---	---	---	---	---	---	---
Participation in Class	-.666[2]	---	-.557[1]	-.543[1]	-.638[2]	-.563[1]	-.776[2]	-.723[2]	-.671[2]	---	-.659[2]
Interest	-.719[2]	---	-.541[1]	-.555[1]	-.671[2]	---	-.561[1]	-.585[1]	-.599[1]	-.586[1]	-.620[1]
GAIN											
Manipulation (N = 12)											
Ability	---	-.649[1]	---	-.699[2]	---	---	---	---	-.667[2]	---	-.611[1]
Productivity & Application to Work	---	-.593[1]	---	-.685[2]	---	---	---	---	-.657[1]	---	-.593[1]
Cooperation	---	---	-.680[2]	-.663[2]	---	---	---	---	---	---	---
Lecture-Demonstration (N = 14)											
Ability	-.609[1]	---	---	---	-.689[2]	---	-.759[2]	-.633[1]	-.540[1]	---	---
Cooperation	---	---	---	---	---	---	---	---	---	---	---
Participation in Class	-.621[1]	---	---	---	-.647[1]	---	-.688[2]	-.608[1]	-.574[1]	---	-.535[1]
Interest	-.602[1]	---	---	---	-.561[2]	---	-.553[1]	---	---	---	---

[1] $p < .05$
[2] $p < .01$

posttest measures. On the other hand, significant correlations in the lecture-demonstration group for the Participation in Class and Interest variables were evident for both improvement and posttest scores. These behaviors in the lecture-demonstration group seem to have a strong relationship to performance on the evaluation instrument following teaching. Only one classroom behavior variable (Interest) in the lecture-demonstration class related to the verbal reasons score on posttest or improvement scores.

D. .iscussion

What is apparent from the results of this comparison of parallel exploratory-manipulative and lecture-demonstration presentations of curricula is that both were equally effective in teaching the simple concepts of electricity to these special class students, at least as measured by the particular evaluation procedure employed.

The failure to confirm the major hypothesis of the study -- that the students in the manipulation group would learn significantly more than those in the lecture-demonstration group is provocative. One possible explanation is that the very small numbers of students involved in this comparison study may not have permitted the differences between presentation to be evident. Educational comparisons or treatments involve gross interventions with enormous uncontrolled sources of variability - namely, the individual child, leaving the other sources aside. Too few. students were involved in this comparison of teaching interventions to override this compelling source of variance.

A second explanation is that the lecture-demonstration unit contained several features that helped EMRs to learn well. Though very different in character from similarly learning promoting features of the manipulation unit, they may be equally effective. The structured form of the lessons may better position the subject matter for the EMRs and hence facilitate their learning of concepts; the drill work may provide good reinforcement for acquired knowledge. That structure -- introduction of material and placing it in context, development, demonstration and conclusion -- aids real learning is an old fashioned concept but perhaps a valid one when one considers the EMR's lack of confidence in learning situations and the frequently avoidant inclination he must cope with while learning. The certainty of knowing what is to be learned, where it is going, and what the purpose of the class is, may help relieve his anxiety towards learning and allow him to be more receptive towards acquiring new knowledge. In addition, the structured mode of the lecture-demonstration unit is akin to the type of presentation the EMR usually receives in other classes in traditional urban schools. Since his concept of how classes should be is probably quite rigid by junior high school, he can probably more easily adjust to a new classroom and settle down more quickly to the business of learning when the class resembles those he is used to. (See <u>Kitchen Physics</u>, Appendix A, for a discussion of EMRs' lack of adjustment to a very loosely structured classroom environment).

Drill work in the lecture-demonstration class may aid the cognitive process by clarifying and reinforcing partially mastered concepts. The attention span of special class students is notoriously short and it is reasonable to assume that many simple but novel concepts cannot be acquired in one presentation alone. Thus, drill exercises may play a primary role in conveying ideas as well as in enriching the topic presentation in the context of a subject matter the student regards as interesting.

If structure and drill work are valuable features in teaching EMRs, why have traditional classes, which show both these characteristics, failed in teaching EMRs over the years? The answer lies in two major characteristics of the electricity lecture-demonstration unit which are not shared by the majority of traditional special class curricula: 1) a deemphasis of the skill of reading and frequent substitution of pictures and diagrams for words, and 2) a well developed, and well constructed assembly of material in which interest and relevance for the children was a predominant concern.

The manipulation class, of course, was much more unstructured than the lecture-demonstration unit, drilled far less extensively, and above all, was conducted in an atmosphere in which the expectation and requirement -- e.g. "make a short circuit, now try different ways to make it"-- are foreign to the manner in which student learning usually occurs. The losses in learning that may result from the undoubtedly difficult adjustment to this atmosphere, however, seem to be more than compensated for by the positive teaching factors of the unit. To reiterate these qualities (which prompted the selection of a manipulation curriculum to test the educational hypotheses of the main study), learning by doing, the main teaching premise of the unit, is an efficient manner for teaching those EMRs who show high nonverbal reasoning aptitude (high learning potential status) relative to their verbal ability. Again, the novelty of the emphasis on materials and exploration probably intrigues many students who are bored with other approaches to learning. The manipulation unit has positive features of its own that enhance the learning process for EMRs.

In summary, the lack of confirmation of the main hypothesis of the study may be due to the fact that both the teaching methods have features that result in effective learning by EMRs. The manipulation unit offers a teaching method that suits the natural learning style of many of the children, as well as an interesting approach. The lecture-demonstration unit gives students good orientation into the purposes of the classes, a framework in which to fit their new knowledge and clarification of this knowledge through exercises. Because of the presence of different but equivalently good educational factors in both electricity units, the students of the two classes may have been able to absorb the same amount of material from the different teaching interventions. Hence, there was little difference in the posttest and improvement scores of the two groups in the pilot study done with the two curricula.

Another contribution to the lack of treatment effect in the study may be the emphasis on the point system for good questions and answers in the lecture-demonstration group. The motivation supplied by the reinforcement device may have greatly increased learning. Informal observation suggests that the two students who showed the highest improvement scores in the lecture-demonstration course were greatly influenced by the point system in their desire to concentrate and work in class. It may be that the very frequent use of this teaching aid may have resulted in higher scores in the class as a whole by involving the students more completely than usual, rendering fair comparison with the scores of the manipulation class questionable.

A final explanation of the failure to confirm the first hypothesis is the generally high quality of both teachers. It may very well be that a good teacher transcends any material he or she teaches, and by enthusiasm, special devices and individual explanations is able to convey subject matter well, regardless of the form of the curriculum. Unfortunately the confounding of teachers and curricula in this study prevents any conclusive statement about teachers in the same way that it interferes with any interpretation about the results of the treatment effect.

The second major hypothesis of the study, that learning potential would determine success in the manipulation unit but not in the lecture-demonstration unit was also not confirmed by the results. The lack of significant interactions of treatment and learning potential on the posttest and improvement scores, and the strong learning potential effect on these same measures tells us that learning potential may very well be a determinant of performance in both these types of units. This unexpected result is also intriguing since it suggests that learning potential may be applicable to a broader range of learning behaviors than has been recognized. It was generally assumed that learning potential taps a nonverbal type reasoning and learning ability that operates best in performance problems like the Kohs Block Designs and Raven's Progressive Matrices, which allow the subject to reason with cues other than verbally acquired knowledge. Thus, it is logical that learning potential be a determinant of a manipulation unit which teaches electricity primarily by having children make circuits and explore problems with their own equipment and minimizes lecturing and discussion as well as reading in conveying information. But its similar role in a unit depending more on skills of listening, discussing and analyzing problems points to a wider role for learning potential which may includes verbal comprehension and performance skills.

E. Comparison of Main Study Results with this (1968) Comparison of Teaching Methods.

A final set of questions is raised by the comparison between the results of the concluding study, comparing teaching methods, with that part of the main study which dealt with EMRs, comparing teaching (with the manipulation method) and with the no instruction condition. The questions were as follows:

1. Are the three samples (main study (1967) nontaught EMRs; main study (1967) taught EMRs; 1968 study EMRs taught (both lecture-demonstration and manipulation groups pooled)) comparable in initial level of knowledge of electricity?

2. Did teaching bring about approximately the same improvement in knowledge in the two studies?

3. Was the learning potential status variable of equivalent importance in the two studies?

Precise answers can be obtained to these questions through the partitioning of the two degrees of freedom (df) for the three treatment groups into one df comparing the main study taught groups with the 1968 study taught groups (A_1), and the other degree of freedom which compares both taught groups with the nontaught group (A_2). Similarly the two df associated with learning potential can be divided into one df associated with the linear component of the continuum (B_1); the other df tests for the nonlinearity of this continuum and compares gainers with nongainers and high scorers (B_2).

1. Comparability of the three groups in initial level of knowledge.

Table 35 presents the F-ratios for the pretest scores on the three samples. Generally, the results indicate that the three samples were comparable in initial level of knowledge. The one significant F-ratio for A_1 on total score, indicates that the 1968 group knew slightly more about electricity than the main study taught group prior to teaching. The one significant multivariate F-ratio is for the linear component (B_1) of learning potential status. This indicates that high scorers knew more about electricity than gainers who in turn knew more than non-gainers.

2. Comparison between the groups in final level of knowledge and in amount learned.

Tables 36 and 37 present F-ratios for the same effects after teaching. First, the difference between the main study (1967) and the 1968 taught groups has disappeared completely. (Posttest multivariate $\underline{F}_m$ = 1.19, N.S.) Nor do the 1967 and 1968 taught

Table 35. Summary of F-Ratios Comparing Pretest Scores of Taught and Nontaught Special Class Students in Major Study (1967) and Pilot Comparison (1968).†

SOURCE OF VARIANCE	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Total	Multi-variate F-Ratio
A. Taught(T)-Nontaught(NT)	2	.62	.93	.98	1.79	.37	1.22	1.36	1.85	3.10	.68
A_1(T'68 vs. T'67)	1	.90	1.77	1.27	3.55	.37	2.38	2.40	3.58	6.13^{2}	.74
A_2(T'67+'68 vs. NT)	1	.35	.09	.68	.03	.37	.07	.32	.12	.08	.62
B. Learning Potential (LP)	2	1.75	2.02	2.44	1.34	1.11	5.54	1.69	5.72	5.86	1.66
B_1(linear*)	1	3.47	3.55	4.87^{1}	2.50	2.16	10.33^{2}	3.37	11.11^{3}	11.38^{3}	2.83^{2}
B_2(quadratic*)	1	.03	.48	.02	.19	.06	.76	.01	.33	.35	.48
A x B	4	2.79	.65	.73	.58	.62	.59	.41	.36	.53	.92
A_1B_1	1	4.00^{1}	.16	.01	1.65	.50	1.22	.36	.44	1.39	.88
A_1B_2	1	3.98^{1}	.02	1.01	.24	.02	1.20	.03	.53	.06	.86
A_2B_1	1	1.55	.43	1.90	.21	.10	.01	1.15	.45	.58	.91
A_2B_2	1	1.64	2.00	.01	.23	1.85	.28	.09	.00	.08	1.05

$^{1} = p < .05$
$^{2} = p < .01$

*For LP linear component, groups were ordered as high scorers, gainers, nongainers. For quadratic component, gainers were compared with the other two groups.

†Significant F-Ratios are Reported only for Linear and Quadratic Effects.

Table 36. Summary of F-Ratios Comparing Posttest Scores of Taught and Nontaught Special Class Students in Major Study (1967) and Pilot Comparison (1968).†

SOURCE OF VARIANCE	df	Simple Circuits	Comples Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Total	Multi-variate F-Ratio
A. Taught(T)-Nontaught(NT)	2	10.72	3.24	9.97	8.93	16.28	4.44	5.72	11.20	11.37	3.05
A_1(T'68 vs. T'67)	1	2.07	.56	.93	.03	.08	.43	2.29	1.00	.51	1.19
A_2(T'67+'68 vs. NT)	1	19.38^3	5.93^1	19.01^3	17.83^3	32.49^3	8.44^2	9.14^2	21.39^3	22.23^3	4.92^3
B. Learning Potential (LP)	2	10.54	7.77	15.88	17.70	10.30	9.36	18.61	18.26	21.79	3.50
B_1(linear*)	1	20.94^3	15.29^3	31.50^3	34.78^3	20.58^3	18.69^3	37.20^3	36.17^3	43.53^3	6.15^3
B_2(quadratic*)	1	.15	.15	.26	.62	.03	.02	.02	.34	.05	.85
A x B	4	1.96	1.35	1.24	1.65	2.26	.75	.91	1.85	1.94	.97
A_1B_1	1	3.39	.17	2.64	.72	.09	.40	.17	.00	.14	.93
A_1B_2	1	.92	.90	.42	.14	.00	.23	.00	.28	.01	.58
A_2B_1	1	3.42	4.30^1	.93	4.95^1	8.85^2	2.32	3.45	7.04^2	7.40^2	1.21
A_2B_2	1	.11	.02	.98	.78	.08	.04	.03	.09	.22	1.16

$^1 = p < .05$
$^2 = p < .01$
$^3 = p < .001$

*For linear component, groups were ordered as high scorers, gainers, nongainers. For quadratic component, gainers were compared with the other two groups.

†Significant F-Ratios are Reported only for Linear and Quadratic Effects.

Table 37. Summary of F-Ratios Comparing Gain Scores of Taught and Nontaught Special Class Students in Major Study (1967) and Pilot Comparison (1968).†

SOURCE OF VARIANCE	df	Simple Circuits	Complex Circuits I	Complex Circuits II	Subtotal I	Simple Circuits Diagrams	Complex Circuits Diagrams	Schematic Diagrams	Subtotal II	Total	Multi-variate F-Ratio
A. Taught(T)-Nontaught(NT)	2	9.19	1.24	7.44	7.88	11.37	3.70	.01	8.16	10.32	22.12
A_1(T'68 vs. T'67)	1	3.80	.01	.01	.95	.01	.29	.00	.00	.06	1.29
A_2(T'67+'68 vs. NT)	1	14.58^{3}	2.47	14.86^{3}	14.80^{3}	22.73^{3}	7.11^{2}	.01	16.32^{3}	20.58^{3}	4.29^{3}
B. Learning Potential (LP)	2	15.12	3.40	3.25	11.63	6.02	1.25	.92	6.80	12.38	2.61
B_1(linear*)	1	30.18^{3}	5.58^{1}	6.43^{1}	22.39^{3}	10.97^{3}	2.02	1.81	13.46^{3}	24.70^{3}	4.77^{3}
B_2(quadratic*)	1	.07	1.23	.07	.87	1.07	.47	.04	.14	.05	.45
A x B	4	1.76	1.35	1.62	1.53	2.48	1.31	1.53	2.40	2.78	1.28
A_1B_1	1	.27	.38	1.27	.04	.05	2.02	.25	.34	.08	.62
A_1B_2	1	4.58^{1}	.52	1.51	.01	.66	.15	.15	.17	.13	1.51
A_2B_1	1	.98	4.47^{1}	3.03	5.86^{1}	9.12^{2}	3.04	5.24^{1}	9.05^{2}	10.61^{2}	1.68
A_2B_2	1	1.22	.03	.66	.21	.07	.02	.47	.02	.28	1.33

1 = $p < .05$
2 = $p < .01$
3 = $p < .001$

*For linear component, groups were ordered as high scorers, gainers, nongainers. For quadratic component, gainers were compared with the other two groups.

†Significant F-Ratios are Reported only for Linear and Quadratic Effects.

groups differ in amount learned (gain score $\underline{F}_m = 1.29$, N.S.). By contrast, whereas the taught groups were comparable to the nontaught group in initial level of knowledge, the taught vs. nontaught contrast (A_2) is highly significant ($\underline{F}_m = 4.92$, $p<.001$), for absolute level of knowledge after teaching, and ($\underline{F}_m = 4.29$, $p<.001$) for amount learned (gain scores). These results are even more impressive when the univariate $\underline{F}$s are inspected.

Learning potential status was an even more important determinant of knowledge of electricity on the posttest than on the pretest. Whereas the multivariate F-ratio for the linear component of LP status was 2.83 ($p<.01$) on the pretest, it was 6.15 ($p<.001$) on the posttest, and 4.77($p<.001$) for amount learned. That the increment in the learning potential effect (linear) was at least partially associated with the teaching intervention and not simply a practice effect is shown by the five significant F-ratios for the interaction of teaching effect, quadratic, and learning potential effect, linear (A_2B_1), on the posttest and the six significant F-ratios for A_2B_1) on gain scores, representing amount learned.

There were no significant effects for the quadratic component of learning potential which compared high scorers and nongainers with gainers.

VII CONCLUSIONS AND IMPLICATIONS

The conclusions of the study are evident and have been repeatedly stated in this report. The high learning potential (high scorer and gainer) special class student who is diagnosed as educable mentally retarded on the basis of a low scholastic aptitude score (e.g., Binet IQ) performed similarly to educationally retarded CA peers on this special nonverbal curriculum. By contrast, the nongainer special class child did not learn from the special curriculum. In terms of functioning level, the nongainer special class child must be considered operating at a mentally retarded level, since he has failed to profit from the teaching experience. This failure to learn is evident and defines the mental retardation in functional terms. While no data is available, it is unlikely that mental age controls (about 10 years) would similarly learn so little. The implication of the nongainers' failure is that there may be an intrinsic defect, rather than a slower rate of development.

These conclusions seem highly warranted from the results of this study. They are more impressive because they appear so consistently with so few subjects per group. Human performance data, such as classroom learning, usually exhibit great variability across subjects and ordinarily would require larger samples unless the effect of the variables are strong and consistent. Because of this expectation of greater variability, the original study projected larger samples which the consistent findings made unnecessary.

The experimental support for the hypothesis underlying the learning potential assessment procedure raises intriguing questions, re: the capabilities of the severely low achieving child from poor socio-economic circumstances. This child's poor school performance has usually been ascribed correctly to his deficiencies in reading, and the language arts more generally. Thus, as the history of school failure and the low scholastic aptitude score, (e.g., Binet IQ), indicate, this child is a poor risk for successfully completing the ordinary school curriculum. When the scholastic aptitude score falls below 75-85 IQ, depending on the operative state law, the child is deemed to be seriously limited in his ability, i.e., intelligence more generally, and is placed in special classes for the retarded which are intended to be special learning situations.

Unfortunate and unintended results come from this strategy. The strategy is intended to facilitate the child's progress by placing him in a more individualized program of instruction geared to his personal requirements and so protect the child from continued failure. There is consistent agreement that special class placements are not educationally facilitating for the child. Controlled studies indicate these children simply do not learn in these classes though they are said to feel more comfortable. One apparent explanation is that the IQ score and the placement tend to result in a self fulfilling prophecy. The teacher does not want to push a child who

cannot learn, assumes he cannot learn, and so does not challenge him sufficiently. And the child does not learn. The child feels the stigma of being placed in a class for dumb children, and ceases to want to, or try to learn (because he has been told he can't?). Also he experiences a certain degree of social stigmatizing by peers and adults. The combined effects result in a terminal placement because the child cannot return to regular grades after any substantial length of time in special class since he has not mastered the skills or materials to do the work except at the lower grade levels, since the teacher and the educational system, as well as psychologists, believe that the low Binet IQ justifies a judgment that the child is incapable of being educated beyond the mental age predicted from this indicator of rate of development.

We would argue that the low Binet IQ score does indicate high risk for successful completion in the ordinary academic school. The question that remains is whether the child with a low scholastic aptitude score and a record of scholastic failure is slow to learn or profit from his experiences, or mainly from the range of experiences available in academic subjects. Given that they are poor students, are they "dumb" in solving reasoning problems more generally? The learning potential assessment tests this latter question by the strategy of presenting nonverbal reasoning problems and allowing the child the option of an individual tutorial in which to learn how to solve that type of problem. The hypothesis is that the child with ability in the learning potential situation will display this ability with other nonverbal kinds of tasks, though not with verbal problems. Thus, the child _is_ intelligent, if we must use that term, but not for school types of subject matters that involve language arts related skills or contents. With subject matters, such as the laboratory science course presented in the present study, the ability evident on the learning potential task predicted performance following teaching reaffirming that this ability is not task-specific and can be made relevant to educational learning situations as well as on experimental learning and reasoning tasks.

It may be that early identification of the prospective school failure who is able by a learning potential criterion, and early intervention to assure him individualized instruction tailored to minimize failure and maximize successful learning, may even alter the poor prognosis indicated by the low scholastic aptitude score (Binet IQ). Given successful and meaningful learning experiences with nonverbal subject matters such as laboratory science or mathematics it may be possible to alter the teacher's view of the child fostered by the low Binet IQ as slow or unable and so continue the process of challenging the child. Also, these continuing experiences of successful learning will continue to keep the child working positively in school and may even embolden him to try harder with language arts materials. This procedure might be facilitated considerably if the language arts materials were suitably tailored for the child.

The findings of the present study supporting the learning potential argument also may help one better understand the seemingly anomalous but consistent evidence that a large proportion of school age special class EMRs attain an independent social and economic adult adjustment. Further, when group differences are compared, there are few significant differences between them and CA peers from similarly poor social backgrounds. Occupationally, then, the poor school age prognosis is not confirmed. The few differences in job level and job security may be due to the stigmatic consequences for the employee of having been labelled retarded, and the fact of these persons' functional illiteracy at the time they leave school.

What becomes evident in the context of the learning potential studies, and which has been supported by the present study, is that these children do have ability to learn and reason with nonverbal problems. It must be this ability which was displayed by the special class high scorers and gainers in the electricity curriculum. It must also be this ability, which becomes manifest on jobs in the years after leaving school, leading, in part, to the controversy re: "pseudofeeblemindedness". It should be apparent that the learning potential argument developed in the present research program can rather easily account for this phenomenon. That is, the seeming discrepancy between the child's performance in school, and the ability he demonstrates after leaving school, is not a discrepancy. Rather, the ability the school age EMR displays after leaving school reflects abilities which the school has not tapped because they are not dependent on adequate verbal expressive or verbal conceptual skills, which are the skills which the school is basically concerned with.

Several educational implications follow from these results:

A. Since certain EMRs show themselves to be very capable when taught with nonverbal exploratory-manipulative material it may be that more such units, in the field of science, mathematics, and/or social studies should be adopted for this type of learning. Further research is of course needed on whether the manipulation and consequent first hand experience with the materials was the crucial factor in expediting learning in the electricity unit, or whether the novelty of the subject matter, its relevance to the student's interest and the fact of the minimization of reading were more important. Such research has practical implications as well as a theoretical one, since supplying materials to each student for these types of units is an extremely expensive undertaking for a school system.

B. Just as high scorer and gainer students with IQs in the educable mentally retarded ranges may require novel presentations and interventions by which to learn that will tend to minimize the negative effects of their poor literacy skills, so, too, would low achieving regular class children with dull to average IQs seem to be able to profit from similar types of specially designed classroom interventions. If the implications of the learning potential hypothesis are correct, then skill in literacy, however small, and the consequent ability to profit, however minimally,

from the reading laden curricula favored in current schools may be the only factor which distinguishes the regular class school failure from the gainer and high scorer. Yet if the low achiever can learn considerably from non-verbal presentations as was demonstrated in the present study, he should be given the opportunity to work with these types of curricula so he can be meaningfully re-engaged in school learning. Again, research is needed with the low achiever to determine which type of curriculum he actually does profit best from.

3. Gainers and high scorers might do better academically in a regular class placement than in a special class. If (a) certain changes were made in curricula to compensate for their reading deficits and/or (b) they were given extensive work in reading to try and bring their skills up to regular class standards. It could be that if gainers and high scorer EMRs were not in special class, with its accompanying stigma of being "dumb", or unable, or slow to learn, they might be as capable of learning and applying new concepts as the regular class students. This regular class placement, with appropriate modifications that would individualize learning opportunities for these high risk students, might result in considerable alleviation in their school difficulties. These changes might require a higher proportion of nonverbal learning opportunities such as this laboratory science unit, and others like it in science, mathematics, and social studies. It would also require considerable expenditure of effort in remedial work in language arts subjects to determine whether these students can improve their competence markedly. For these severe school failures, it might also require explorations of the utility of providing tangible rewards or incentives to further motivate them to work in the language arts subjects in which they have experienced considerable failure. It may be that within the context of successful work in the nonverbal academic presentations and tangible rewards for trying to achieve more satisfactorily with the negatively loaded language arts materials, these students' motivation to work productively in school might be increased considerably, and result in markedly higher levels of achievement. However, a regular class placement for the non-reader based on a learning potential rating in which he would be left to fend for himself without special attention, would only do him a disservice, and continue the debilitating pattern of failure.

It may be, however, that the "true" reason for the pattern of severe school failure and the low scholastic aptitude score which resulted in the special class placement is a disability in efficiently processing verbal-conceptual materials. If there are suspicions of this type of disability, intensive efforts must be expended to determine whether it is an intrinsic deficiency or a function of poor prior experience at home and in school. If these intensive efforts do not yield an appreciable upgrading of skills, then this failure may point to the major deficit or defect which may be typical of the child who is classified as educable mentally retarded. Alternate programs of educating these able but nonverbal children would then have to be formulated which would direct themselves toward maximizing these children's strengths in the nonverbal areas, and seek to have

them obtain a maximal level of proficiency in the verbal-conceptual areas, such as reading. It should be self-evident that success in any of these types of determinations will be heightened by intervening at younger chronological ages than those of the students participating in this study so that the negatively reinforcing cycle of failure might be short circuited before it becomes established.

VIII. REFERENCES

American Association for the Advancement of Science. Commission on Science Education. SCIENCE, A PROCESS APPROACH. Experimental edition. Washington, D.C., 1963, Parts 1-4.

American Association for the Advancement of Science and the Teaching Center, University of Maryland. THIRD REPORT OF THE INFORMATION CLEARINGHOUSE ON NEW SCIENCE AND MATHEMATICS CURRICULA. College Park, Maryland, March, 1965.

Atkins, J.M. SOME EVALUATION PROBLEMS IN A COURSE CONTENT IMPROVEMENT PROJECT. Journal of Research in Science Teaching, 1963, 1, 129-132.

BATTERIES AND BULBS. Trial Teaching edition, Elementary Science Study of Educational Services, Inc., Watertown, Mass., 1966.

Bloom, B. (Ed.) TAXONOMY OF EDUCATION OBJECTIVES. Handbook I, Cognitive Domain. New York: David McKay Co., Inc., 1956.

Budoff, M. LEARNING POTENTIAL AMONG THE EDUCABLE RETARDED. Progress Report, 1965, mimeo'd.

Budoff, M. LEARNING POTENTIAL AMONG INSTITUTIONALIZED YOUNG ADULT RETARDATES. American Journal of Mental Deficiency, 1967, 72, 404-411.

Budoff, M. SOCIAL AND PSYCHOMETRIC TEST DATA CORRELATES OF LEARNING POTENTIAL STATUS, 1968, dittoed.

Budoff, M. & Friedman, M. LEARNING POTENTIAL AS AN ASSESSMENT APPROACH TO THE ADOLESCENT MENTALLY RETARDED. Journal of Consulting Psychology, 1964, 28, 434-439.

Budoff, M., & Pagell, W. LEARNING POTENTIAL AND RIGIDITY IN THE ADOLESCENT MENTALLY RETARDED. Journal of Abnormal Psychology, 1968, in press.

Cattell, R.B. A CULTURE-FREE INTELLIGENCE TEST. Journal of Educational Psychology, 1940, 31, 161-179.

Davis, A. & Eells, K.W. DAVIS-EELLS GAMES: DAVIS-EELLS TEST OF GENERAL INTELLIGENCE OR PROBLEM-SOLVING ABILITY, MANUAL. Yonkers-on-Hudson, N.Y.: World Book, 1953.

Giddings, M.G. SCIENCE FOR THE DISADVANTAGED. Teachers College Record, 1966, 67, 435-442

REFERENCES

Goldstein, L., & Sheerer, M. ABSTRACT AND CONCRETE BEHAVIOR. Psychological Monographs: General and Applied, 1941, 53 (whole No. 239).

Goodlad, J.I., Von Stoephasius, Renata, & Klein, M.F. THE CHANGING SCHOOL CURRICULUM. The Fund for the Advancement of Education. New York, 1966.

Hagen, Elizabeth. Unpublished paper presented at American Psychological Association meetings, 1966.

Harrison, R.H., & Budoff, M. A FACTOR ANALYSIS OF THE LAURELTON SELF CONCEPT AND BIALER'S LOCUS OF CONTROL SCALES. 1968 mimeo'd.

Hedges, W.D. TESTING AND EVALUATION FOR THE SCIENCES IN THE SECONDARY SCHOOL. Belmont, California: Wadsworth Publishing Company, Inc., 1966.

Kirk, S.A. In The Slow Learner, Segregation or Integration. Edited by F.J. Schonell, J. McLeod and R.G. Cochrane. St. Lucia: University of Queensland Press, 1962, 235-240.

KITCHEN PHYSICS. Trial Teaching edition, Elementary Science Study of Educational Services, Inc., Houghton Mifflin, 1965.

Kohs, S.C. INTELLIGENCE MEASUREMENT. New York: Macmillan, 1923.

Lisonbee, L. TEACHING SCIENCE TO THE DISADVANTAGED PUPIL. The Science Teacher, 1963, 30, 18-21.

Lombard, J.W. PREPARING BETTER CLASSROOM TESTS. The Science Teacher, 1965, 32, 33-38.

McFall, R.W. THE DEVELOPMENT AND VALIDATION OF AN ACHIEVEMENT TEST FOR MEASURING HIGHER LEVEL COGNITIVE PROCESSES IN GENERAL SCIENCE. Journal of Experimental Education. 1964, 33, 103-106.

Malkin, S. THE CULTURALLY DEPRIVED CHILD AND SCIENCE. Science and Children, 1964, 1, No. 7, 5-7.

Nelson, H. EVALUATION OF OBJECTIVES OF SCIENCE TEACHING. Science Education, 1959, 43, 20-27.

Olsen, J. THE VERBAL ABILITY OF THE CULTURALLY DIFFERENT. Educational Forum, 1965, 29, 280-284.

Reiner, W.B. MEETING THE CHALLENGE OF RECENT DEVELOPMENTS IN SCIENCE AND ASSESSING THE OBJECTIVES OF MODERN INSTRUCTION. School Science and Mathematics, 1966, 66, 335-341.

Riessman, F. THE CULTURALLY DEPRIVED CHILD. New York: Harper and Row, 1962.

Riessman, F. LOW INCOME CULTURE, and ADOLESCENT AND THE SCHOOL. National Association of Secondary School Principals, 1965, 49, No. 300, 45-49.

Rouse, Sue T. EFFECTS OF A TRAINING PROGRAM ON THE PRODUCTIVE THINKING OF EDUCABLE MENTAL RETARDATES. American Journal of Mental Deficiency, 1965, 69, 666-673.

Walbesser, H.H. CURRICULUM EVALUATION BY MEANS OF BEHAVIORAL OBJECTIVES. Journal of Research in Science Teaching, 1963, 1, 296-301.

Williams, W.C. THE PALS TESTS: A TECHNIQUE FOR CHILDREN TO EVALUATE BOTH PARENTS. Journal of Consulting Psychology, 1958, 22, 487-495.

APPENDIX A.

RESULTS OF A PRELIMINARY TRIAL OF THE EX-PERIMENT WITH THE KITCHEN PHYSICS CURRICULUM

APPENDIX A.--RESULTS OF A PRELIMINARY TRIAL OF THE EXPERIMENT WITH THE KITCHEN PHYSICS CURRICULUM

In order to gain some expertise in teaching and evaluating an exploratory-manipulative science unit, and to pilot test the hypothesis that gainers and high scorers would learn more than nongainers, a trial teaching run was conducted in the spring of 1966. The basic unit for the study was Elementary Science Study's Kitchen Physics, a unit concerned with some physical properties of liquids. By experimenting with water, soapy water and cooking oil, and using common household materials, e.g., aluminum foil, paper towels, newspaper, etc., children are introduced to such properties of liquids as viscosity, weight, and surface tension. They also learn processes of science, such as observing, predicting outcomes, and using equipment. There were several reasons for the choice of this unit -- a. it seemed to be the best exploratory unit available at the time, b. terminology and explanation is kept simple, c. the subject matter, "drawn from a child's own environment" (Kitchen Physics Manual, 1966, p. 1), was thought to be particularly appealing to educationally disadvantaged children, and d. the unit was designed for 5th to 8th graders, an appropriate age range for the experimental and control groups of young adolescents to be used in the study.

The evaluation plan used for Kitchen Physics was a laboratory test in which children carried out tasks and commented on what they had done. This format was suggested by the obvious measurable outcomes of Kitchen Physics. Initially the attempt was made to review the unit, its objectives, content, etc., so that an evaluation plan would be oriented toward the kind of information or processes the pupils should have mastered by the end of the unit. A close examination of Kitchen Physics revealed that the major content concepts were tension-adhesion and viscosity. It was noted however, that these were, in reality, very poorly delineated; they could both be summed up in the basic theme of the unit -- water grabs together better than soapy water. Other concepts of the unit such as the various weights, and absorption and evaporation rates of liquids were better defined, but the whole unit did not hang together thematically.

Processes of science, on the other hand, were well developed in Kitchen Physics. Observation skill, prediction from a previous experience, translation of a problem into numbers, the notion of error, and the use of equipment, were all integral parts of the unit. After consultation with ESS staff and examination of the teacher's manual, the set of problems presented in the unit was defined and a laboratory test formulated which emphasized processes. This type of test, administered individually, enabled the children to demonstrate their process skills and knowledge in an active manner, and thus maximize their responses. The following types of items were included in the laboratory test: 1. skill of Observation, 2. ability to offer a Solution or reason behind observed phenomena, 3. Prediction on the basis of previous knowledge gained in the test, 4. Reason for the Prediction given, 5. use of Equipment, 6. use and understanding of a Balance (Balancing), 7. understanding of the equivalence of Weights (related to 6.), and 8. Knowledge of Experimental Error. A description of the test, in which the juxtaposition of content and skill areas can clearly be seen is given below. (Phrases in parentheses indicate the category of response.)

The Test Procedure Described

The student was first asked to place two drops of water on a glass plate with an eye-dropper (use of Equipment), push them together and describe what happened (Observation). He then repeated the performance with soapy water, stated any difference observed between the water and the soapy water (Observation), and hypothesized what would happen with cooking oil in the same series of tasks (Prediction, and Reason for Prediction); he was also asked to offer a reason (Solution) for any difference observed between the water and the soapy water. Next he was presented with a transfer task. To a level medicine cup of water he was asked to add "all the water you can with the eye-dropper until the first drop spills." (Equipment); he then did the same with a level medicine cup of soapy water, previously hypothesizing whether it would heap to the same level, higher or not as high as the water (Prediction and Reason for Prediction). Following this task he was asked to observe the two levels, detect the higher one and offer a solution as to why it was the higher (Observation, Solution). Another transfer task was the prediction of the higher level between fifty drops of water and fifty drops of soapy water, the performance of the experiment, and the offering of an explanation for the observed outcome (Prediction, Equipment, Solution). On both transfer tasks S was asked to predict how oil would behave in comparison to water and soapy water (Prediction and Reason for Prediction).

The examiner then instructed S in the use of a simple balsa wood balance; following this S was presented with one simple balancing problem -- to make the previously used fifty drops of water and fifty drops of soapy water balance -- and one highly complicated problem involving a real understanding of the balance (Balancing, Weights). S was asked to compare the weights of "a teaspoon" of water and "a teaspoon" of oil, using different weighing measures -- small washers, staples or paper clips for each. His prowess in use of the balance, and his understanding of the process of weighing were observed (did he understand that washers were as valid weight units as pounds or ounces?) Finally a clincher question was asked -- which was heavier, the washer or staple or paper clip measure? The best answer, of course, was that it was impossible to tell since one cannot compare two objects weighed with different units. After the Balance and Weights questions, S was shown a chart of different sample results in weighing fifty drops of water; the results were all slightly different and S was asked to give as many reasons as possible for the discrepancies (Knowledge of Experimental Error). Finally four strips of paper, one dipping into a beaker of water, one into oil, one into soapy water, and one into alcohol, were shown to S; he was asked to predict the absorption rate after ten minutes and the absorption - evaporation rate after twenty-four hours (Prediction).

The Sample Described

The trial run of Kitchen Physics was conducted at the Royal E. Robbins School in Waltham, Massachusetts in Spring, 1965. An industrial town located near Boston, Waltham specializes in technical light industry such as watchmaking and refers to itself as "The Precision City". The population of 55,413 (U. S. Census, 1960) is 80% native born but also includes large numbers of Italian and

French Canadian immigres. The Robbins School of the Waltham Public School system is composed entirely of eight special classes; the school principal at the time of this study, Miss Eleanor Maloy, was also director of special education in Waltham. Many children in the Robbins School and in other special classes in the city are French Canadian immigres for whom the language barrier is great enough to force a normal or dull normal IQ into the retarded ranges.

Thirty-four non-brain damaged educable mentally retarded children were used for the trial teaching. Twenty-seven of these were boys and seven were girls. All children came from the Robbins School with the exception of five girls bussed into the Kitchen Physics classrooms from their own special class located in a nearby junior high school. The chronological age range of the children was ten years, ten months to eighteen years, one month. The group was divided into two classes according to age. Although there is a CA overlap (see Table A - I) the two groups will be referred to as the "older" and "younger" classes. The older group was composed of sixteen boys and two girls (N = 18); the younger of eleven boys and five girls (N = 16). All the children had previously received the Kohs learning potential procedure and had been designated as high scorers, gainers, or nongainers. The breakdown for each class by learning potential status was high scorer - 5, gainer - 12, and nongainer - 6, for the younger class. The imbalance by Kohs status was unavoidable and a function of the available sample since there were few high scorers (5) and nongainers (7) and they fell into one or the other group by CA and/or maturity level. Table A - I gives the ranges, means, and standard deviations for CA and IQ for the older and younger groups and then the same measures by learning potential.

Procedure

Following the administration of the Kitchen Physics pretest to all children, the two classes were run twice weekly for nine weeks from March 17th to May 26th, 1966. In both classes the following topic areas were covered: 1. comparing drops of different liquids -- water, soapy water, oil, alcohol, etc., on several surfaces, 2. measuring the beading point of liquids, 3. timing the rate of flow of different liquids from plastic bottles, 4. weighing the liquids (including the principles of balancing), measuring the tensions of different liquids with a tensiometer constructed from a balance, and 5. measuiring the absorption and evaporation rate of liquids on different materials. Though the two groups pursued the same activities, the time proportion spent on each was quite different, making for two distinct curricula. The activities were evenly spread over the eighteen class periods of the older group, but in the younger one all lessons, except those involving the balance, were covered in the first nine class periods. The younger group worked for an extensive four and a half weeks on balancing, problems of weighing, distance of weights from the fulcrum, etc. Another important difference between the curricula was that the older class spent some time in discussion of results obtained in experimenting, with special attention to what each child found and the differences between the findings. The younger group, on the other hand, held almost no discussions about results.

Table A-I. Ranges, Means and S.D.s for CA and IQ in Kitchen Physics Sample

	CA (in Months)	IQ
Younger Group		
Total		
(N = 16) -CA		
(N = 15) -IQ		
Range	130-197	47-93
Means	158.94	69.63
S.D.	18.88	11.29
Older Group		
Total		
(N = 18)		
Range	159-217	58-109
Means	182.4	78.22
S.D.	13.03	11.65
Younger Group		
Gainers		
(N =9)		
Range	139-197	166-217
Means	165.33	185.42
S.D.	21.00	12.21
Younger Group		
Non-Gainers		
(N = 7)		
Range	130-164	47-69
Means	150.71	59.86
S.D.	11.49	7.91
Older Group		
High Scorers		
(N = 5)		
Range	159-195	72-109
Means	177.40	89.40
S.D.	13.36	12.31
Older Group		
Gainers		
(N = 12)		
Range	166-217	67-88
Means	185.42	75.08
S.D.	12.21	5.85

Teaching was done by members of the ESS professional staff on the premise that they could demonstrate Kitchen Physics at its richest. Each group was conducted by one teacher and one teaching assistant. In addition, a member of the project staff was present in each class to observe the children's progress and classroom behavior. The same test was administered individually to each child before the classes began and immediately following their termination.

Results

Table A-II shows the mean pretest, posttest, and gain scores for the four different groups of subjects (older high scorers, older gainers, younger gainers, younger nongainers) on the different category items of the test (Observation, Solution, Prediction, Reason for Prediction, Equipment, Balancing, Weights, and Knowledge of Experimental Error). Since there are an unequal number of items per category, the figures on the table represent mean score per category. An analysis of total gain (t test) scores revealed a significant difference ($p < .001$) between pretest and posttest scores, when the data was combined over all the groups.

A comparison between the total gain scores of the "older" and "younger" groups indicated that the older group increased its score significantly more than the younger group ($p < .01$). This fact can probably be attributed to the older group's relatively longer and more even exposure to the variety of Kitchen Physics topics, as well as to the higher CA. A comparison by learning potential status of the total gain scores of the younger group showed no significant difference between gainers and nongainers. This latter finding, however, must be qualified by the significantly higher score of the gainers on the pretest ($p < .001$), suggesting a possible ceiling effect on the posttest.

A series of within categories comparisons between the gain scores of the older and younger groups revealed: 1. no significant differences in Observation, Prediction, Equipment, and Weights; 2. greater gain of the older group in the Solution ($p < .05$), Reason for Prediction ($p < .01$), and Knowledge of Experimental Error ($p < .05$) categories; and 3. greater gain of the older group on the Balancing questions ($p < .02$). The results (1) showing no significance and the results (3) showing greater gain of the younger group on Balancing must be qualified by the higher pretest scores of the older group in the categories of Observation, Equipment, and Balancing, suggesting the possibility of a ceiling effect on the posttest. The greater gain in Solution, Reason, and Knowledge of Experimental Error in the older group reflects the greater exposure to these topics in the older class, and a greater proficiency in verbal explanation naturally evidenced by the older students. Finally the significantly greater gain of the younger group in Balancing can be directly attributed to the longer period of training on balance problems that this group experienced in class.

Within categories comparisons were made between the younger gainers and nongainers on the two areas of the test they did best on -- Balancing and Weights. The younger group did significantly better than the older group ($p < .05$) on the former category and showed more gain, though not significantly more, than the older group in the latter one. In the comparison

Table A-II. Mean Score per Item on Eight Test Categories for the Four Subject Groups on Kitchen Physics Pretest, Posttest and Gain Scores.

High Scorers - N= 5
Older Gainers - N=12
Younger Gainers - N= 9
Non-Gainers - N= 7

Pre-Test	Observation	Solution	Prediction	Reason	Equipment	Balance	Weights	Chart
High Scorers	2.30	1.50	2.20	1.88	2.78	3.00	2.28	1.75
Older Gainers	2.60	1.70	2.15	1.71	2.71	2.86	2.14	1.92
Younger Gainers	2.52	1.50	2.06	1.48	2.41	2.40	1.70	1.88
Non-Gainers	2.14	1.10	1.80	1.20	2.09	1.93	1.15	1.20
Post-Test								
High Scorers	2.75	2.33	2.85	2.68	3.00	2.94	2.75	2.75
Older Gainers	2.92	2.52	2.66	2.52	2.98	2.96	2.65	3.00
Younger Gainers	2.62	1.64	2.38	1.58	2.47	2.78	2.38	1.76
Non-Gainers	2.26	1.43	2.14	1.40	2.41	2.43	1.92	1.40
Gain Score								
High Scorer	.45	.83	.65	.80	.22	-.06	.47	1.00
Older Gainer	.32	.82	.51	.81	.27	.10	.51	1.08
Younger Gainer	.10	.14	.32	.10	.06	.38	.68	-.12
Non-Gainer	.12	.33	.34	.20	.32	.50	.77	.20

of the gain scores between the learning potential groups within the younger class, however, no significant difference was found between gainers and nongainers on the Balancing and Weights categories (in fact the nongainers demonstrated a greater, though not significantly greater, gain). Again, the possibility of a ceiling effect operating on the scores of the gainer group might be considered.

Discussion

Interpretation of the Kitchen Physics intervention must be qualified by certain shortcomings in the experimental design. Unfortunately, due to the subject population from which the sample was drawn, there were no non-brain damaged older nongainers to compare with the group of older gainers; therefore, all comparisons of gainers and nongainers were necessarily limited to the younger group. Furthermore, no control group was included to check for the effect of having taken the test independent of the effect of having participated in the intervention. Intuitively, one might hypothesize that such a practice effect would be large in a laboratory type of test. This fact makes the dependent variable (gain score) problemmatic and throws suspicion on all significant gain scores. Finally, the classroom routines were not equivalent enough to make older-younger group comparisons valid. This problem can be seen most clearly in the results of the Balancing category, and the Experimental Error question. In the former case, the significant difference of gain scores of the younger group over the older group is obviously due to the extensive training this group received on balancing problems. In the case of the Experimental Error question, the significantly greater gain of the older vs. the younger group is again attributable to the time spent on group discussion of results and differences of experimental findings in the classroom.

A second problem in the foregoing analyses of gain scores is the ceiling effect on posttest scores. The older group started at a consistently higher level than the younger group and obtained near perfect scores on a number of questions in all categories on the posttest. Hence, they obviously did not have an opportunity to demonstrate their increased understanding following exposure to the unit. This ceiling effect is directly attributable to the structure of the test. As a series of performance tasks, it provided many opportunities for learning - a pretest score could often reflect knowledge gained during the test rather than knowledge acquired previous to the test. The ceiling effect was also compounded by the fact that total areas of the test, such as Observation, were simply too easy on the posttest for many of the children.

A final difficulty in the evaluation of learning in the Kitchen Physics intervention lies with the open-endedness of the measuring instrument. Categories such as reason for Prediction, Solution, and even Observation were heavily dependent on the student's ability to explain or willingness to talk at length about the problem. Thus, the evaluation scheme leaned on one of the very measures it had tried to avoid -- verbal ability. The ceiling effect evident for the older group on many of the categories is partially accounted for by their greater verbal facility and greater assurance in communicating

with the examiner. The open-endedness of the scheme also made marking difficult. The instrument did not lend itself easily to an objective scoring system and many inconsistencies were apparent in the rating scheme finally devised.

Conclusion

The Kitchen Physics trial run was not designed as a formal pilot study. Its purposes were to give the project staff familiarity with the unit and an assessment of the unit's worth, some experience with problems of curriculum evaluation and some guidelines in running a "special" special class as well as to pilot test some of the project's hypotheses. Due to the reasons enumerated in the Discussion, it was impossible to conclude anything definitive about the success of the unit with EMRs in general and high scorers and gainers in particular. However, through the experience, prowess was gained in the other areas of know-how -- evaluating, selecting and presenting an exploratory-manipulative unit to disadvantaged EMRs.

The most direct conclusion from the Kitchen Physics experience was the need for a more objective evaluation instrument. The variation in pretest levels, caused by the learning opportunities on the test, pointed up the necessity for a better assessment method of initial knowledge, whereas the "ceiling" performances observed in the verbal categories of the posttest, and the unavoidable inconsistencies in the marking system of the verbal responses, called for a less verbally dependent measure. Finally, the ceiling responses attained in many of the "nonverbal" questions on the posttest underlined the need for a test with a greater range of difficulty. With these considerations in mind, it was decided to design an objective short answer test with a wide range level of difficulty for the experimental science intervention for 1966-67.

The most important conclusion from the trial run, gained informally from observation of the classes, was that Kitchen Physics would not work with the EMR subject pool and would have to be replaced by a science unit that was more acceptable to them. Theoretically, the maturity level of the total group used in the trial teaching was equivalent to the 5th to 8th grade level that Kitchen Physics was designed for, but in actual fact the older students were far too sophisticated for it, and the unit really appealed to neither group. The unit did not represent "science" to the children. The disadvantaged child may respect science, (Riessman, 1962, p. 13) but he has certain well defined notions as to what science is; it may involve machines and scientific instruments, but no plastic bottles and wax paper. The older group felt themselves quite superior to the home-made, unmechanical materials in Kitchen Physics and experimented with them reluctantly; they often refused to experiment with the same equipment two days in a row, feeling they had learned all there was to learn the first time they had worked with it. The younger group was not disdainful of the materials, but did not perceive them as objects for scientific inquiry either. They spent their time in class messing with, rather than experimenting with, the liquids. All in all, indications showed that Kitchen Physics was neither interesting nor acceptable to the students. A new unit was required as a base upon which to demonstrate the hypotheses of the study.

A final realization from the Kitchen Physics experience was that certain recommended classroom practices which worked well with normally achieving middle class students were entirely unsuitable for disadvantaged EMRs. The teacher style demonstrated in both classes was very casual and open-ended. Children were encouraged to explore with the materials they received, but were given very little direction in how to explore; instructional aids to record findings and help tie children's attention to the lesson were also minimal. The materials themselves were meant to be sources of unlimited interest and exploratory ideas. The teaching procedure was quite often a simple presentation of materials -- liquids, eye-droppers, balances, etc., accompanied only by the directions, "see what you can do with these things," or "come get what you need". The lack of structure in classroom teaching also extended to disciplinary techniques. Rules of the classroom were unclear; boundary lines of behavior were not set up and infractions that were quite obvious to the children themselves, such as thievery of materials and dumping water on the floor, were largely ignored by the teachers. The teachers held very rigidly to the maxim, "if they're interested, they'll behave," and breakdown of classroom behavior was often attributed to the failure of the materials to work their natural magic rather than to other causes.

The loose classroom procedure evident in the trial teaching run is a direct outgrowth of the ESS emphasis on creativity and independent exploration and has been used widely and successfully in ESS classrooms. It is directed, however, towards the internally motivated child who seizes upon interesting materials, discovers problems they pose, and tries to solve them by asking questions of his teacher, or by pursuing his own routes of thinking. It is directed towards the "normal" middle class achiever who knows how, and likes to "explore" intellectually, or is so conditioned to responding to the demands of school that he will explore when the teacher asks him to whether or not he is so inclined. His responsiveness to learning demands is paralleled by a sensitivity to behavioral demands whether they are clearly explicated or not. He quickly feels out those behaviors which bring the teacher's approval and those which do not, and stays within the correct boundaries.

The type of student represented in the Kitchen Physics experiment has neither the ability to explore nor the readiness to please the teacher unless he really respects the teacher as someone to reckon with, and not as a "paper" authority figure. The science materials are not of intrinsic interest to him and the worthwhileness of the project must therefore be "sold" to him via his interest in "science" in general, an appeal to the practical advantages of the unit, or some other method. This child does not know how to conduct an independent inquiry with the materials either. In the Robbins experiment, the directive "come and get what you need" was often a signal for a period of apathetic boredom or horseplay rather than directed exploration. The children did not see possibilities for exciting discovery in the materials, and whatever guidance was extended to demonstrate the excitement was too abstract to reach them. Disadvantaged students such as these children require active direction and articulated goals to aim for, (Malkin, 1964, p. 6) which in turn necessitates the breaking down of instructions and problems into concrete steps aided by teaching devices such as charts and diagrams. The world presented to the Robbins students in Kitchen Physics was too open-ended for their grasp. Finally, the disciplinary philosophy of "if

they're interested,"they'll behave," does not go over in a pupil culture which typically regards a serious classroom as one with clear lines of structure and rules of discipline and a serious teacher as one who is not only likeable, but a good classroom manager as well. Such a teacher must demonstrate both his fairness and his ability to handle the "testing" of the children before they can accept him as a leader (Riessman, 1962, p. 83). The teachers in the Kitchen Physics classes were too ready to avoid confronting such "testing" behavior by the students, and so never really won the respect of the students as leaders of the class.

Many of the classroom problems of Kitchen Physics, the children's unwillingness to experiment, the easy boredom and general misbehaviors were encountered in further stages of this project. They were partially the natural results of introducing something completely new to resistant, failure-avoidant learners, and especially of introducing units whose natural style depends on qualities, like initiative and freedom, which are not associated with classroom learning by disadvantaged EMRs. The trial teaching experience enabled the project staff to observe at first hand the children's problems in learning, and helped the research staff plan ways to facilitate the children's acceptance of the new units in subsequent stages. After the Kitchen Physics experience it was decided to increase the structure in future classes, to modify the units used with an eye toward more concrete goals, to provide instructional aids such as notebooks and charts as aids for the child to record what he learns as he proceeds, and finally, to encourage the teachers to firmly direct the classes in learning procedures and behaviors.

The results of the Kitchen Physics trial teaching run, in summary, were 1. decision to select another unit of greater interest to disadvantaged adolescents, 2. decision to evaluate it by a more objective measure with a wider range of difficulty, and 3. decision to provide more structure in the experimental science classroom, both in terms of learning and of behavior. It was felt that with these modifications a good test of the learning potential hypothesis in an educational setting could be obtained.

APPENDIX B.

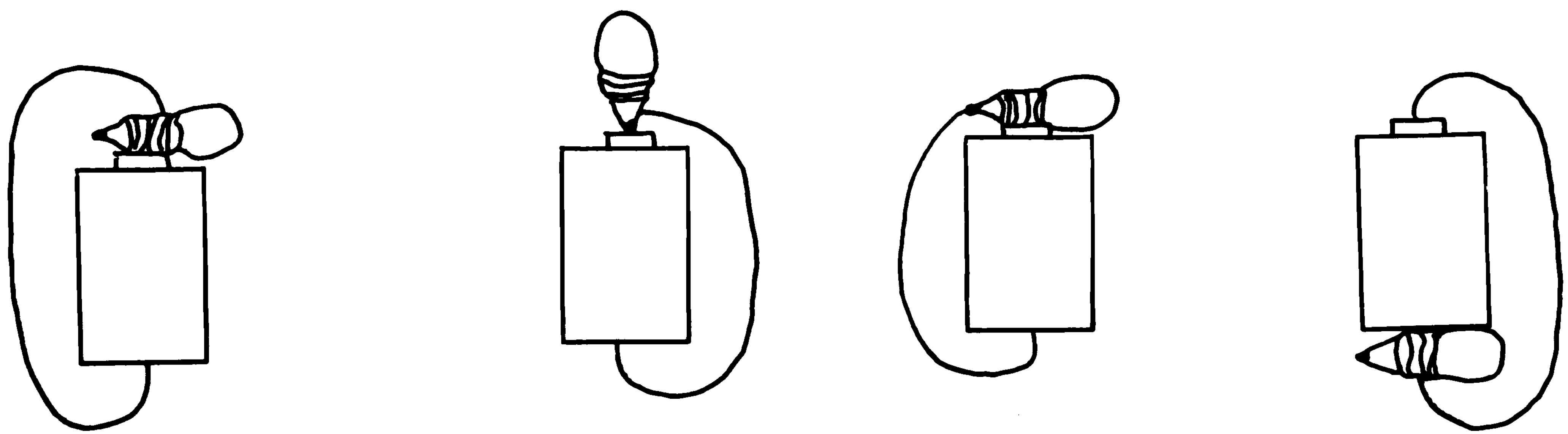
THE DIAGRAMS SECTION OF THE ELECTRICITY EVALUATION INSTRUMENT

____________________ ____________________

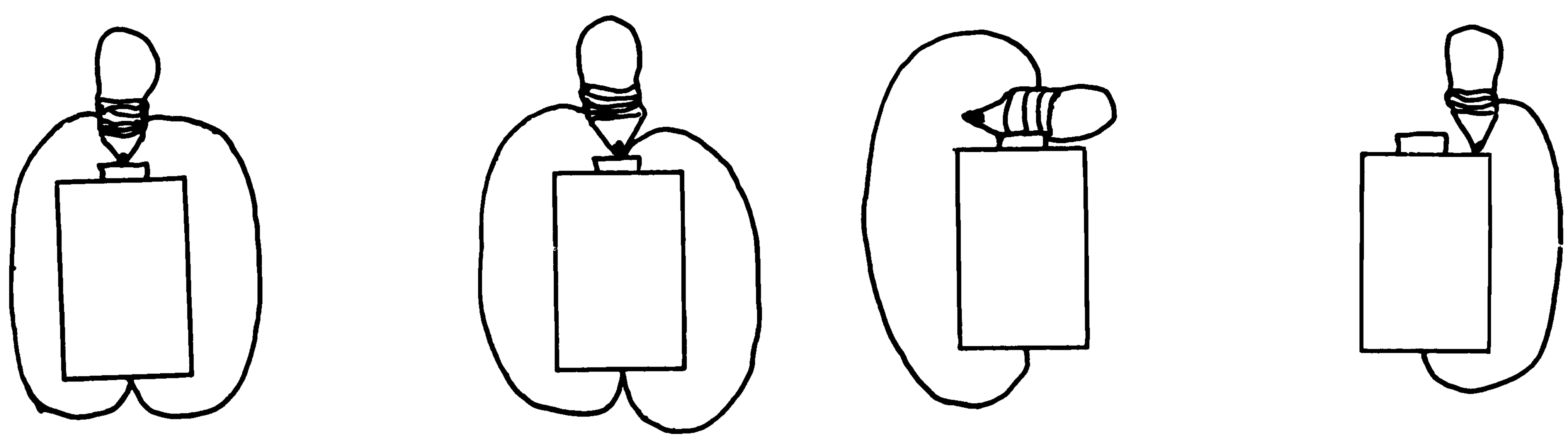
(NAME) (DATE)

PART I: Put an X under the one that lights:

1)

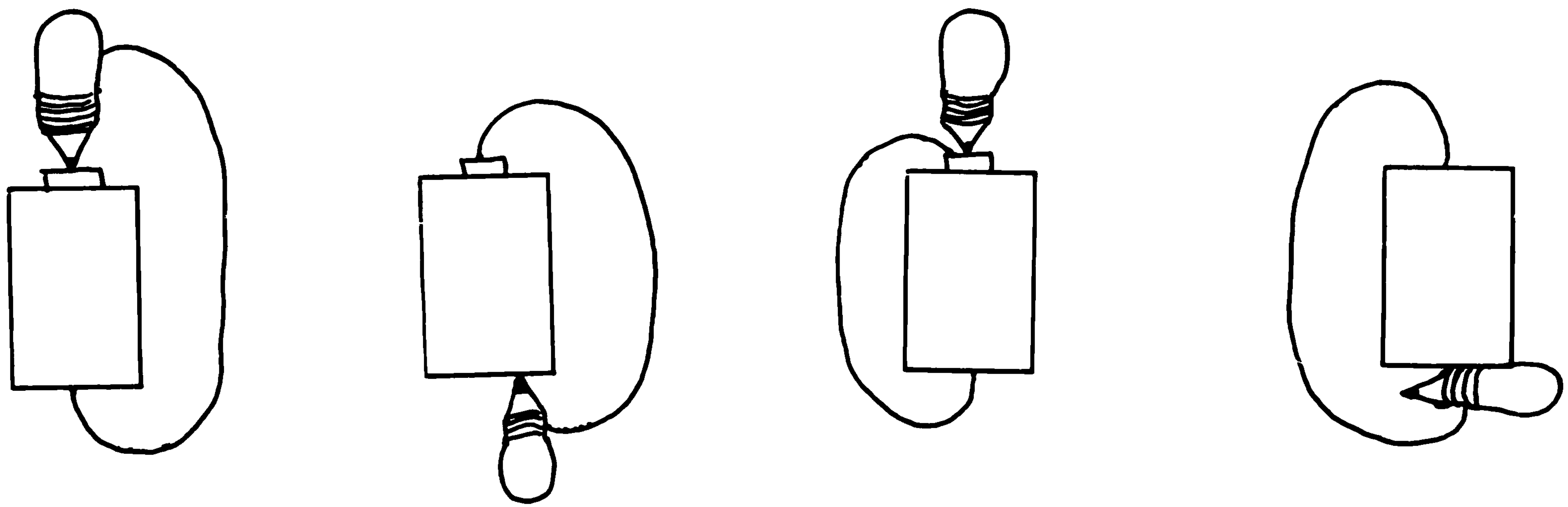


2)

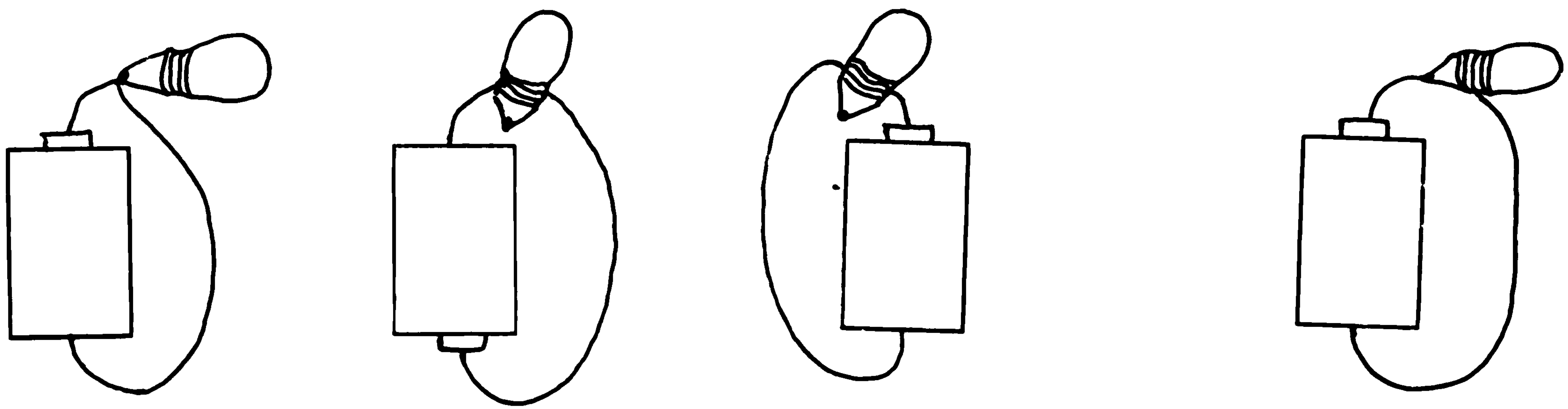


Put an X under the one that lights:

3)

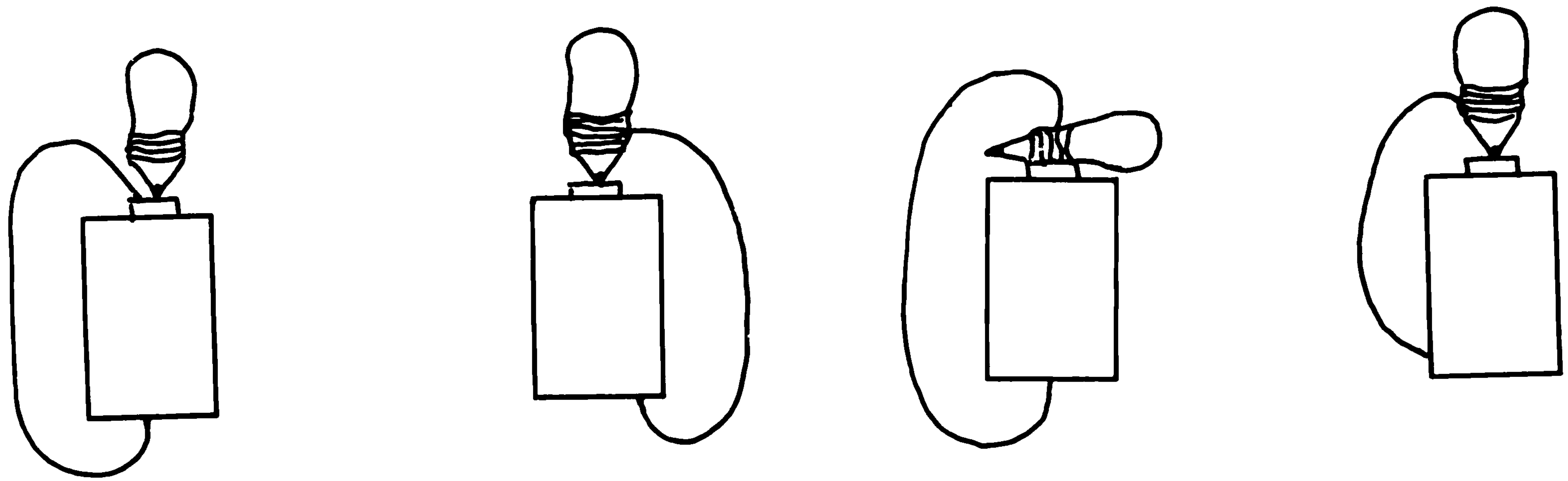


4)

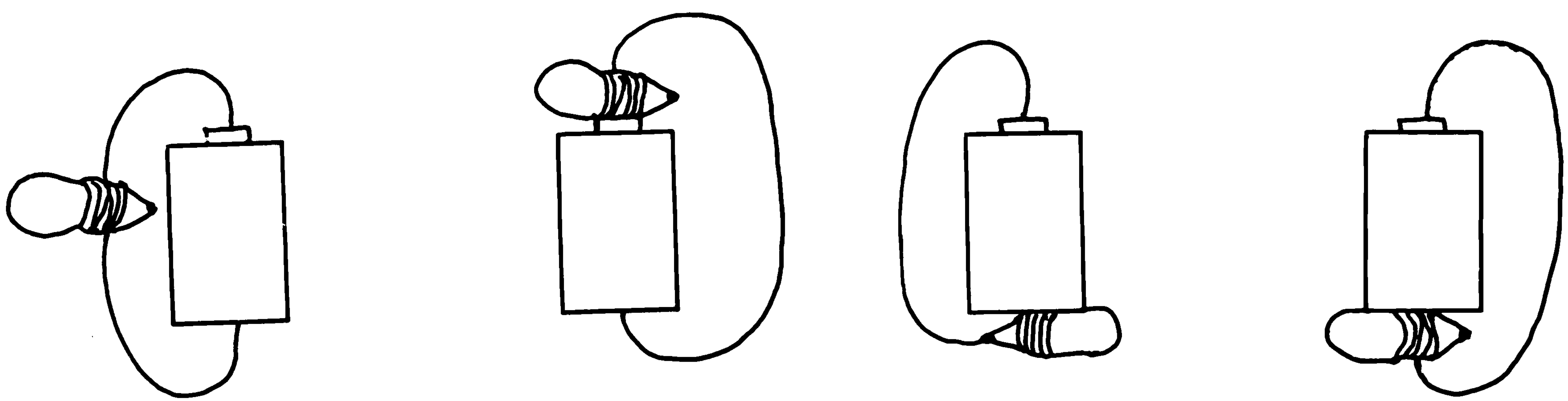


Put an X under the one that lights:

5)

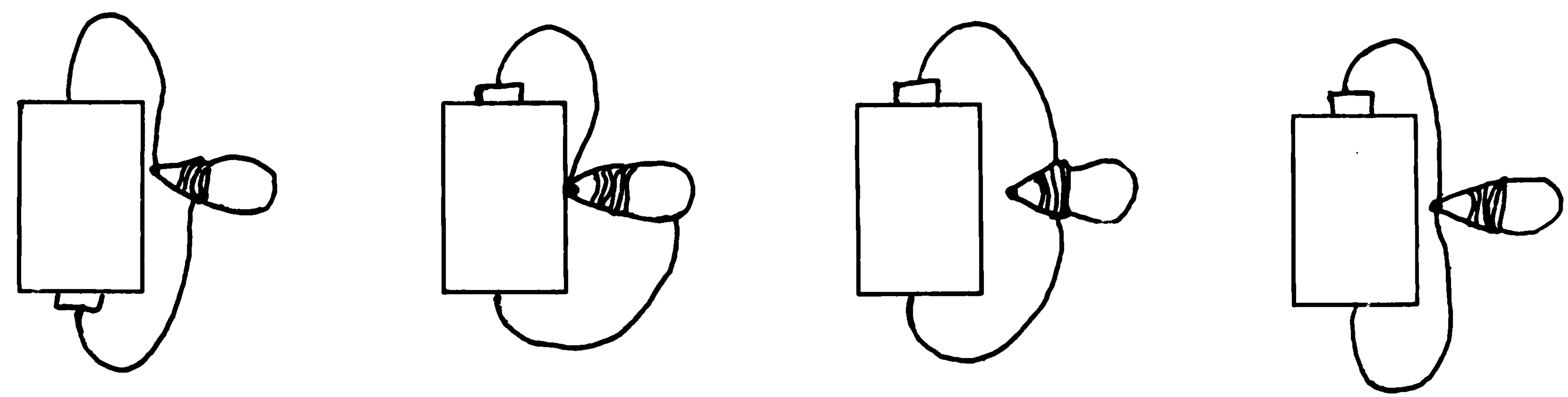


6)

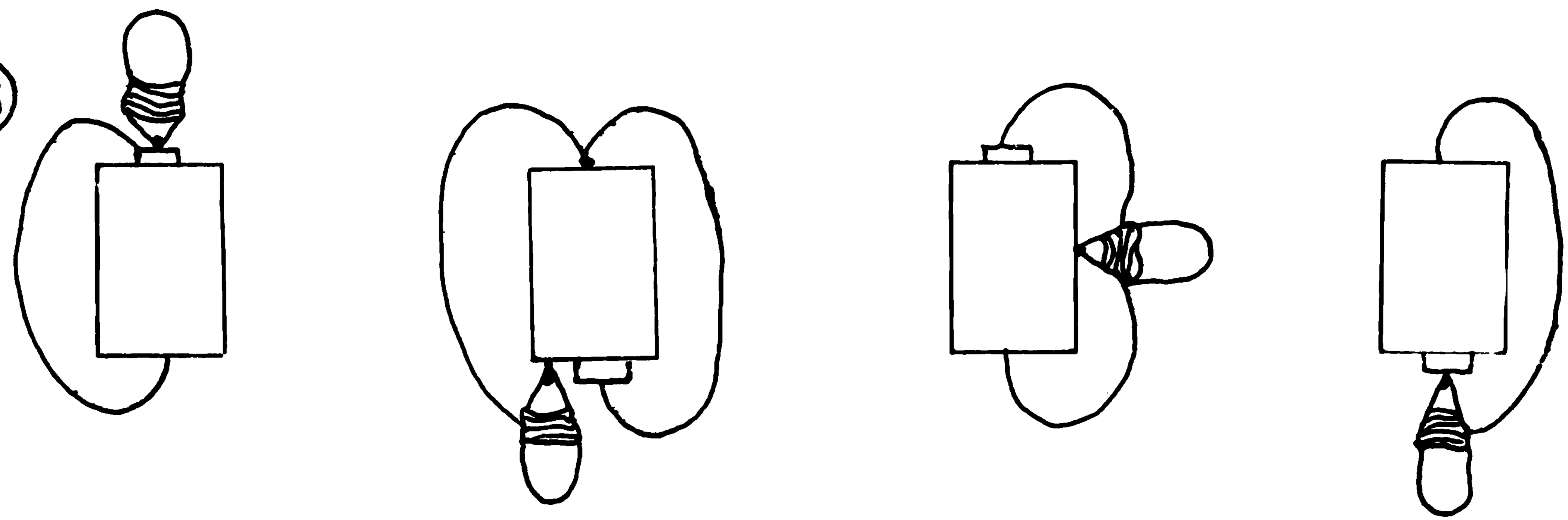


Put an X under the one that lights:

7)

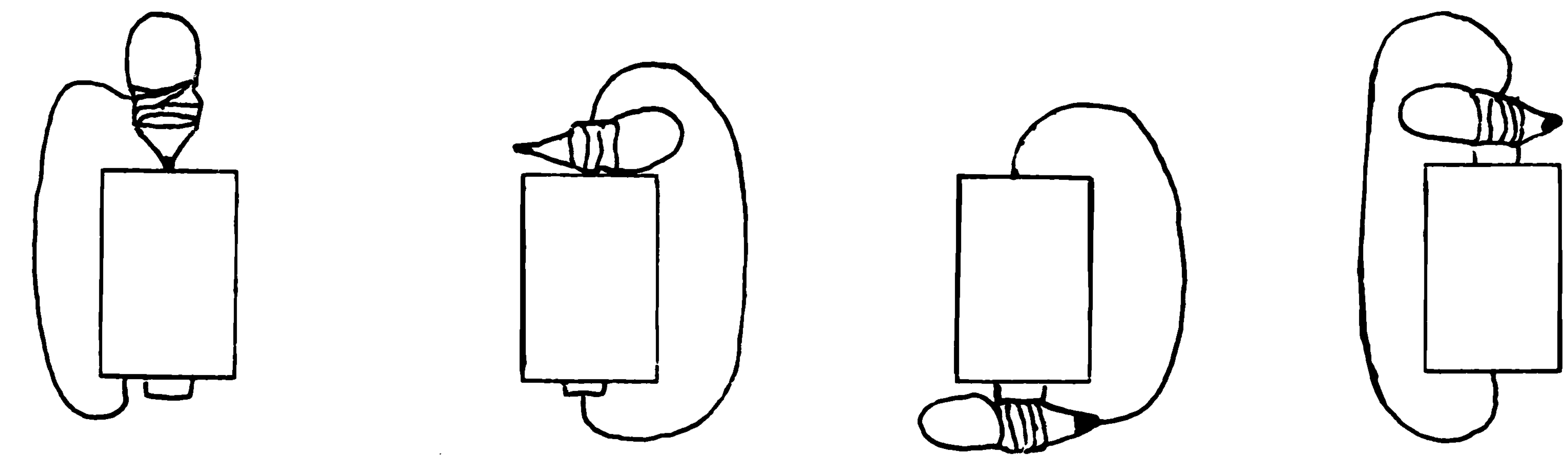


8)

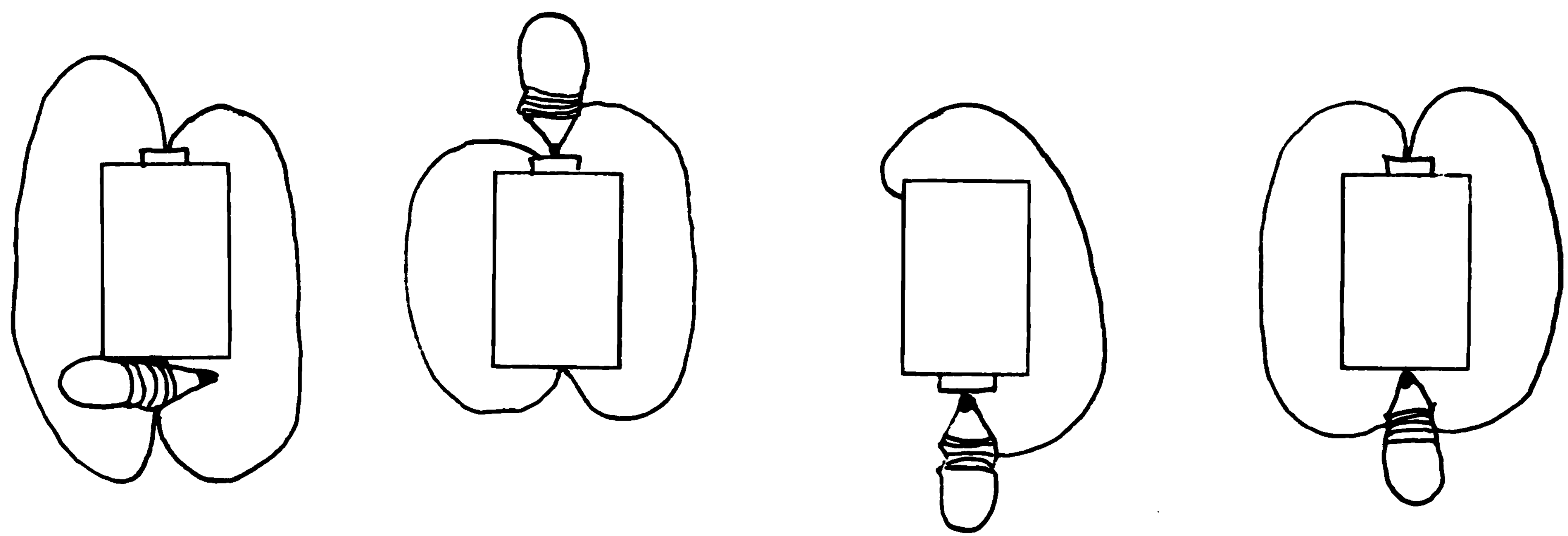


Put an X under the one that lights:

9)

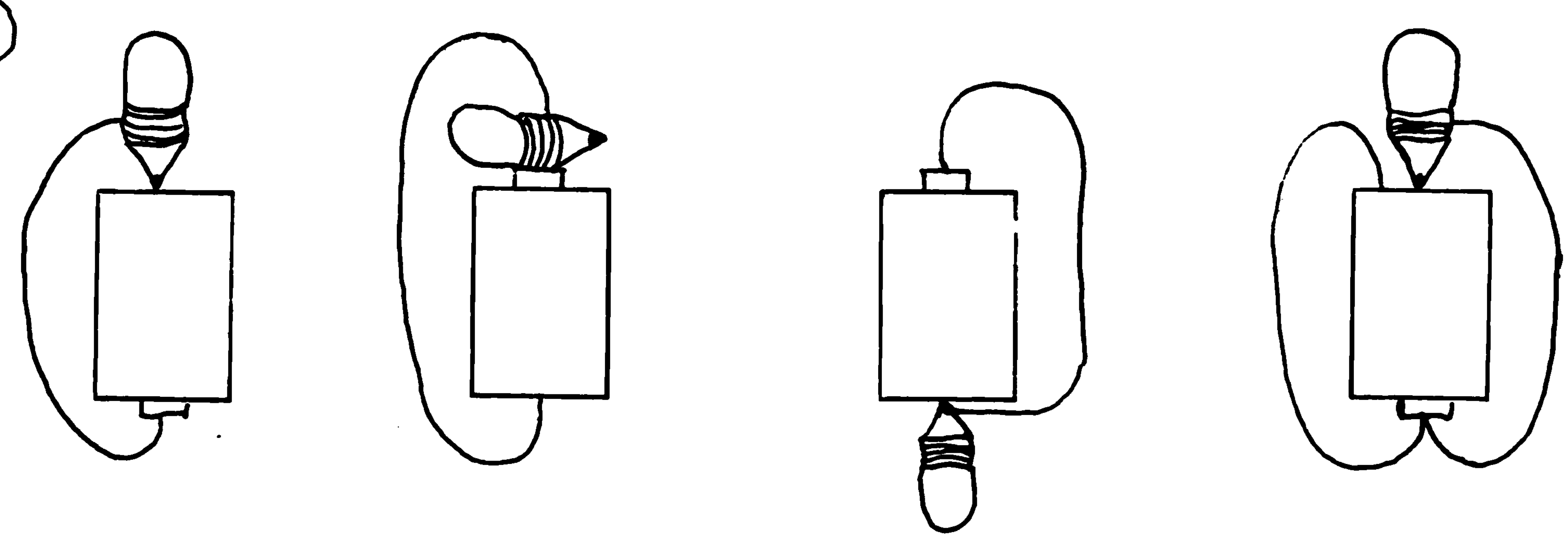


10)

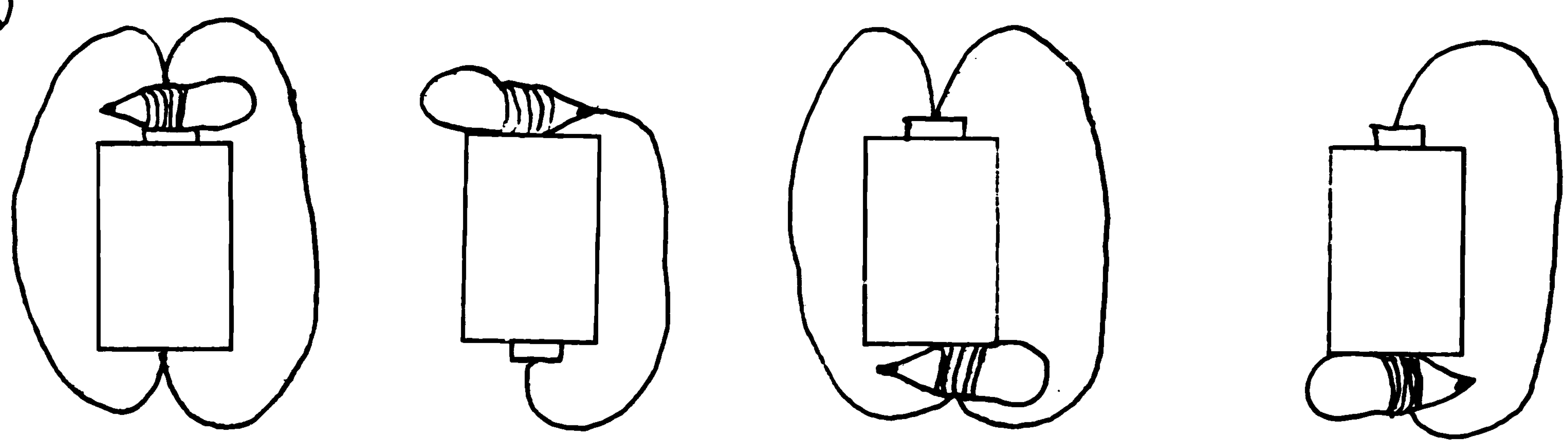


Put an X under the one that lights:

11)



12)



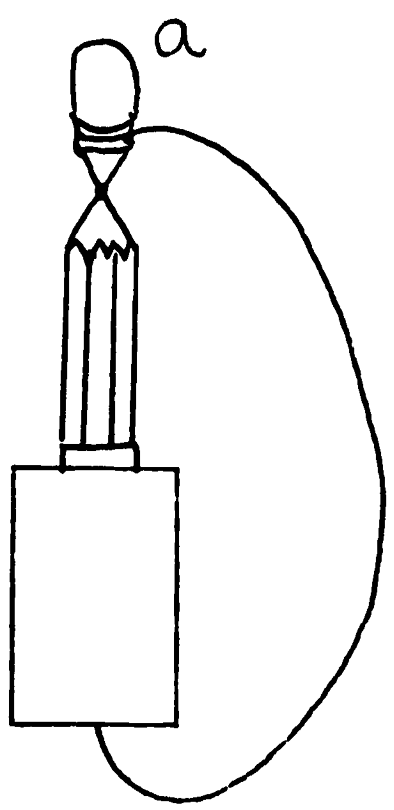
END OF PART I

STOP

PART II:

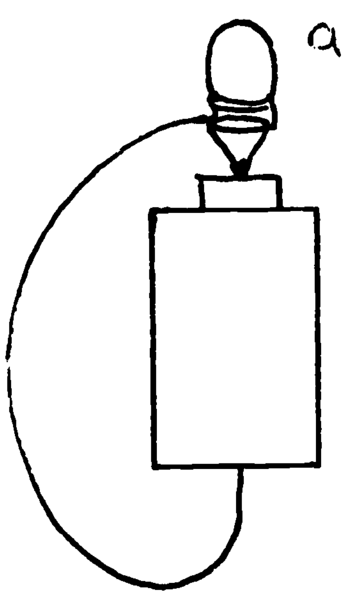
For each set-up, put S for Same (Standard)
B for Brighter
D for Dimmer (or Darker)
N for will Not light

1)



a. ____

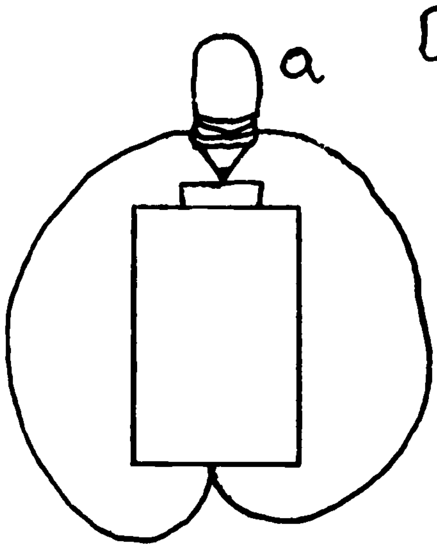
2)



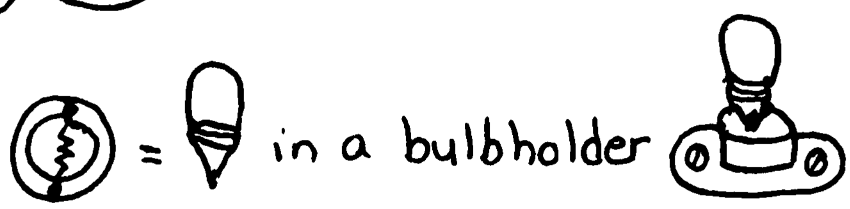
a. ____

S for Same (or Standard)　　D for Dimmer
B for Brighter　　N for Not Light

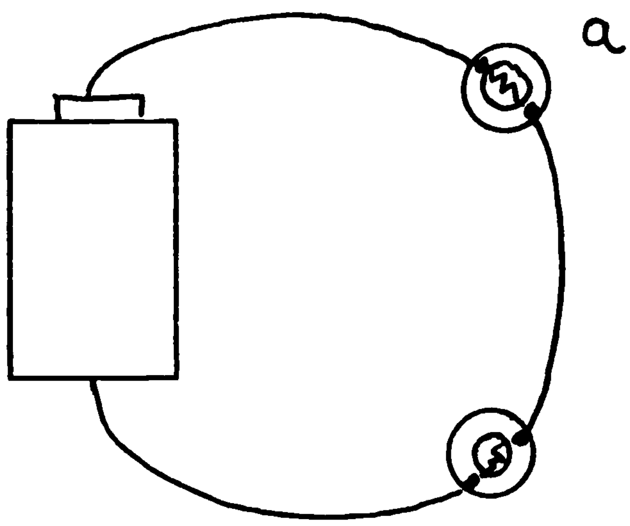
3)



a. ______

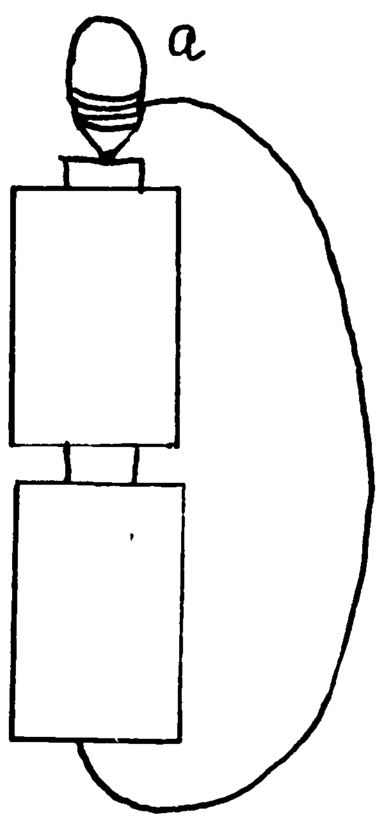


4)



a. ______

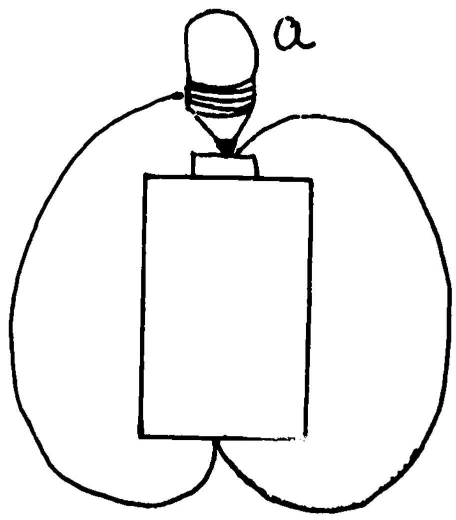
5)



a. ______

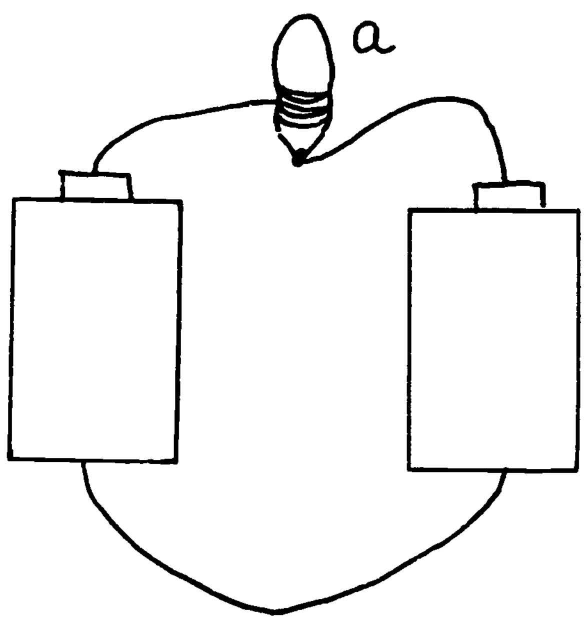
S for Same (or Standard) D for Dimmer
B for Brighter N for Not Light

6)



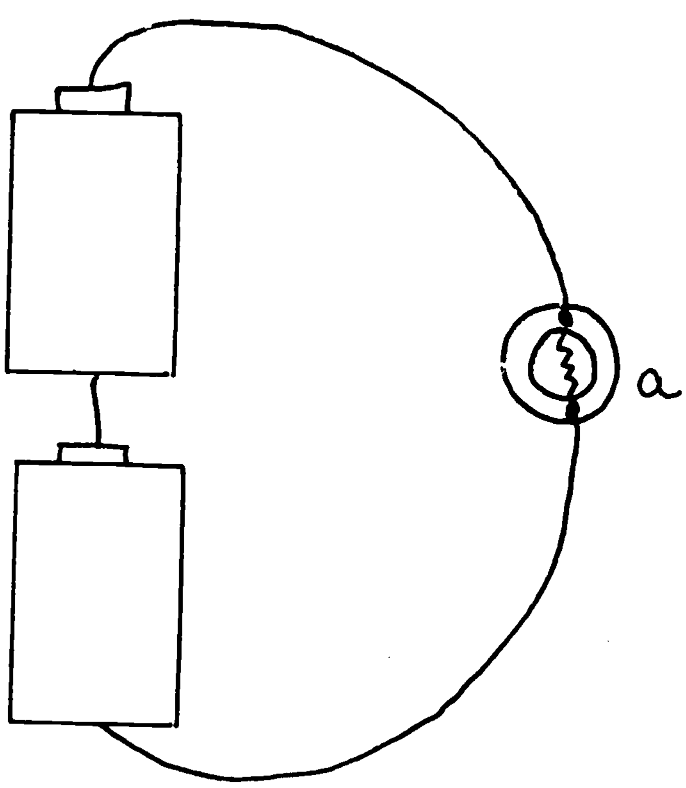
a. ______

7)



a ______

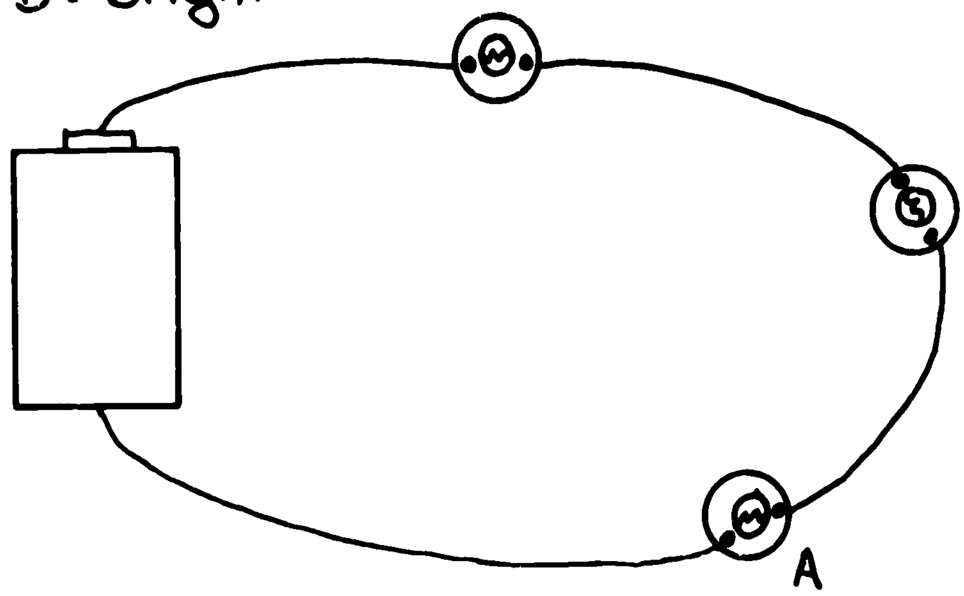
8)



a ______

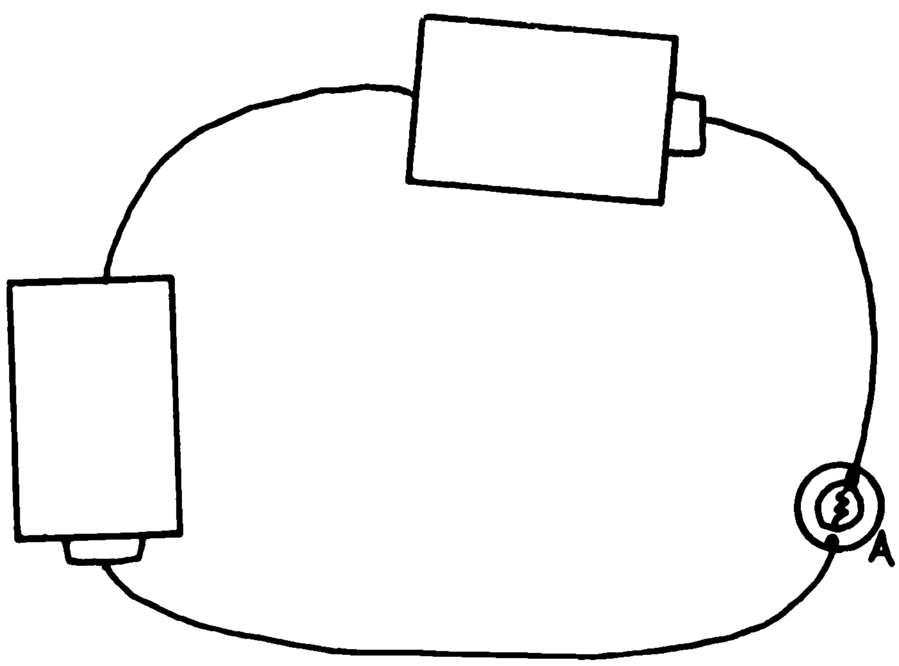
S = Same (or Standard)　　　D = Dimmer
B = Brighter　　　N = Not Light

9)



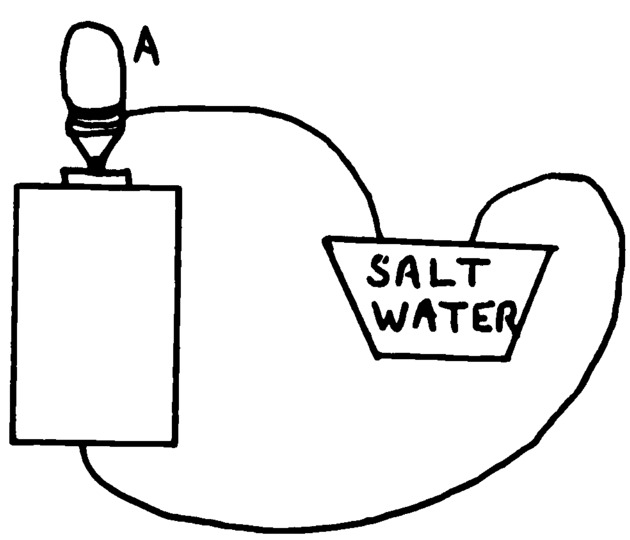
A ______

10)



A ______

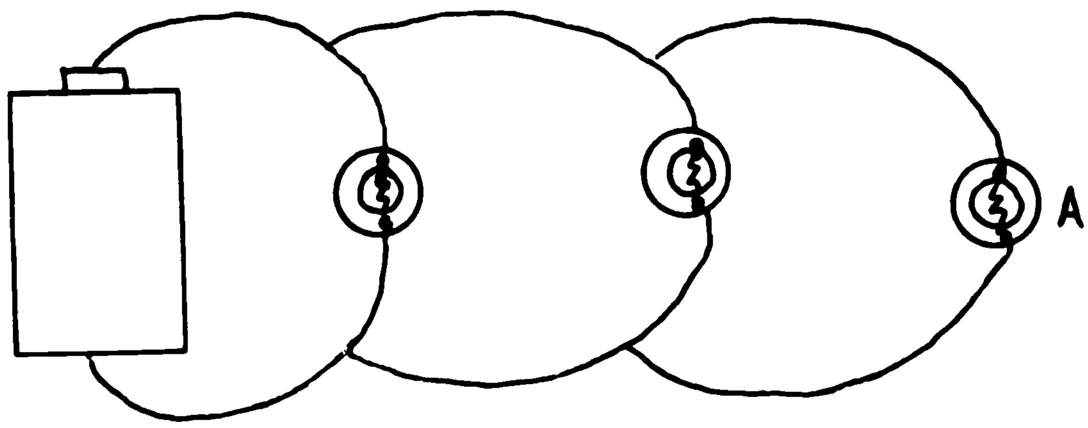
11)



A ______

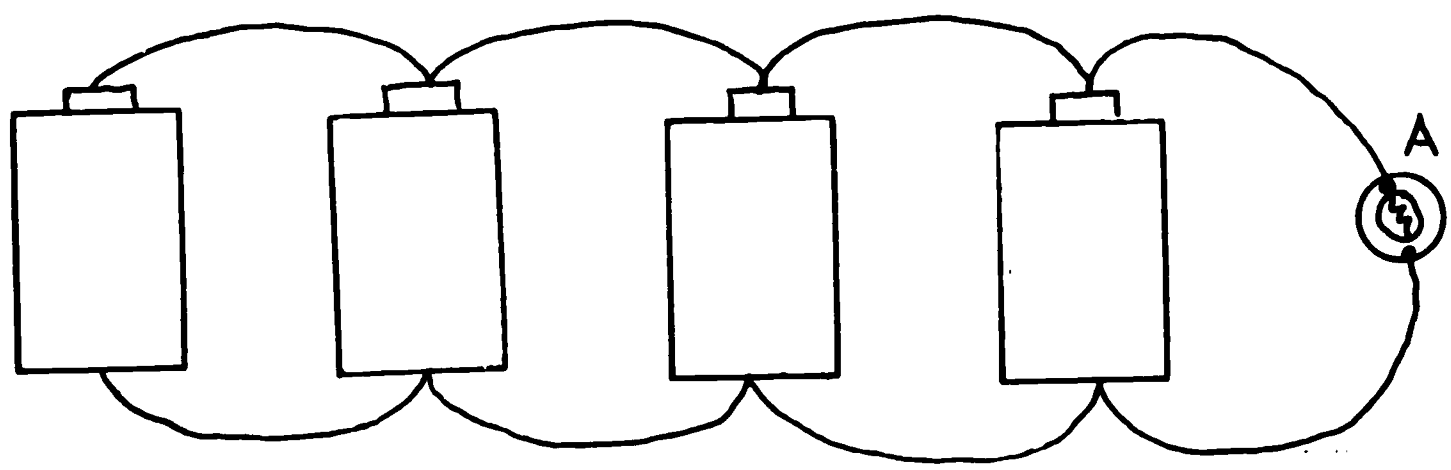
S for Same (or Standard) D for Dimmer
B for Brighter N for Not Light

12)



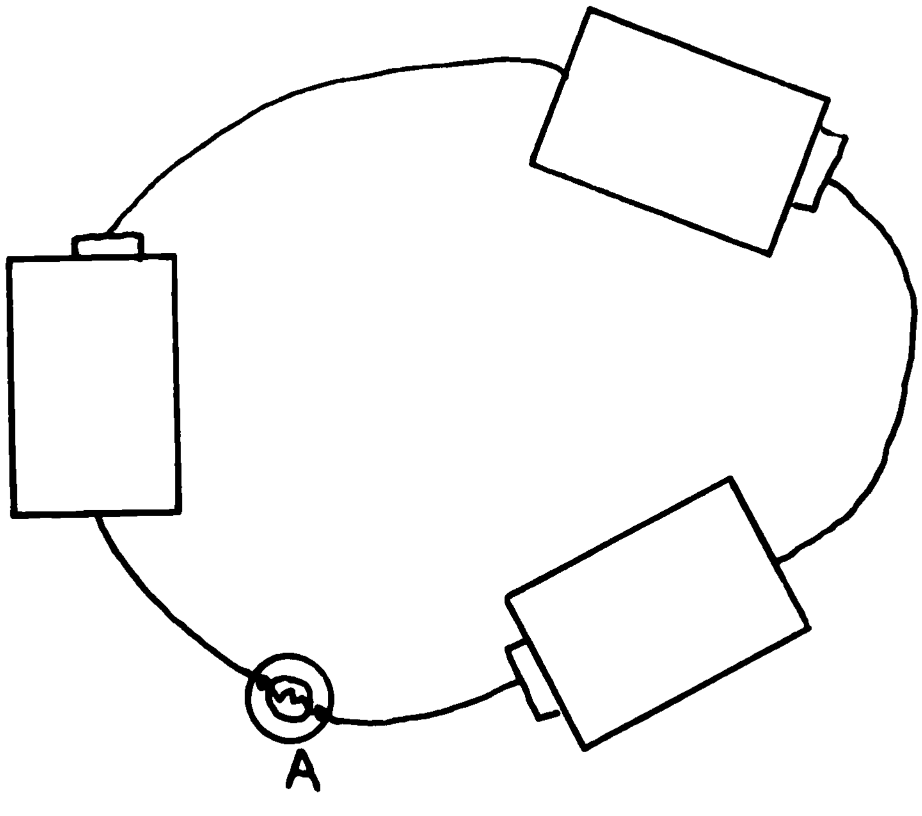
A ______

13)



A ______

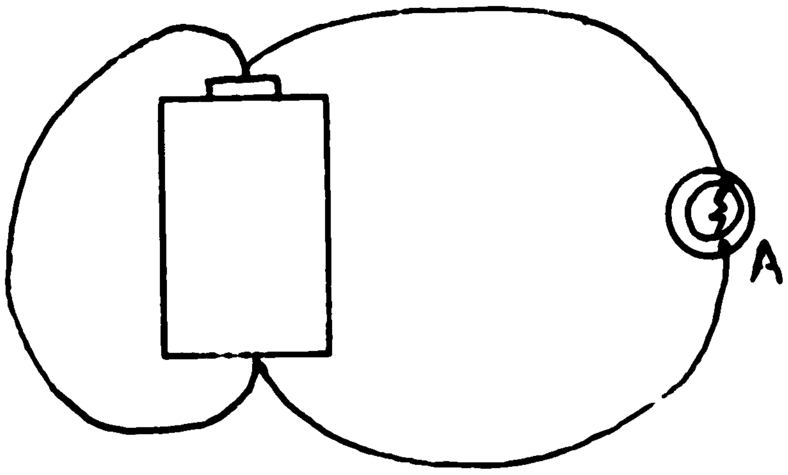
14)



A ______

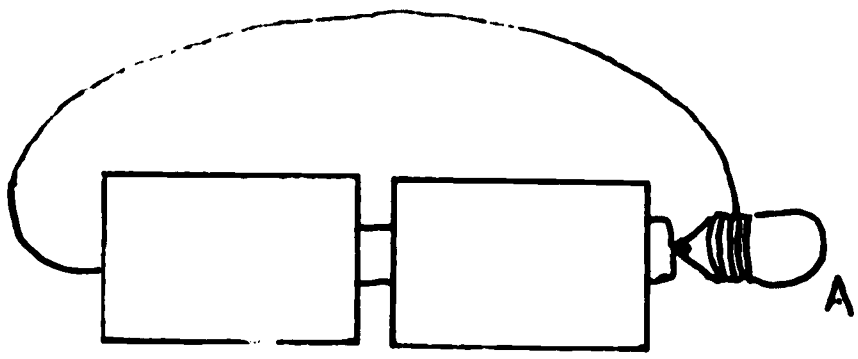
S for Same (or Standard)   D for Dimmer
B for Brighter   N for Not Light

(15)



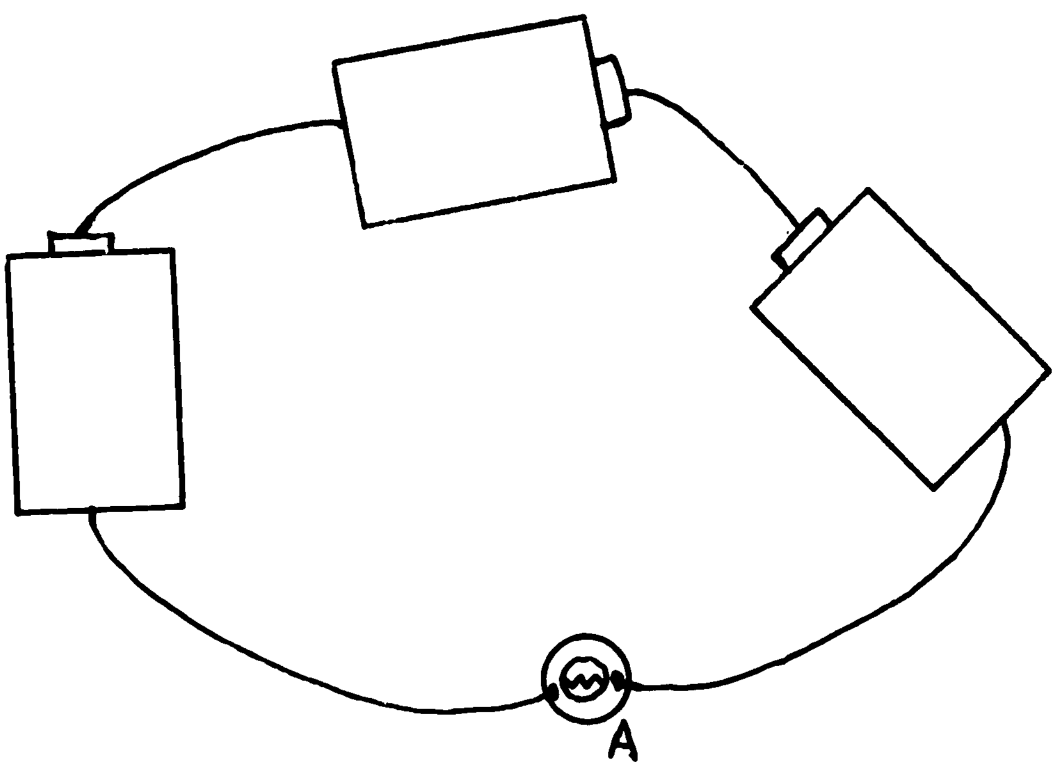
A ________

(16)



A ________

(17)

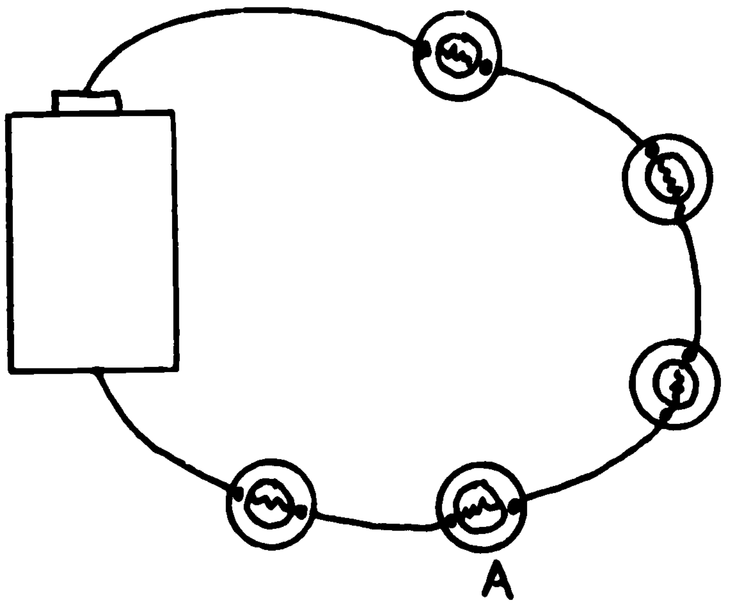


A ________

S for Same (or Standard)
B for Brighter

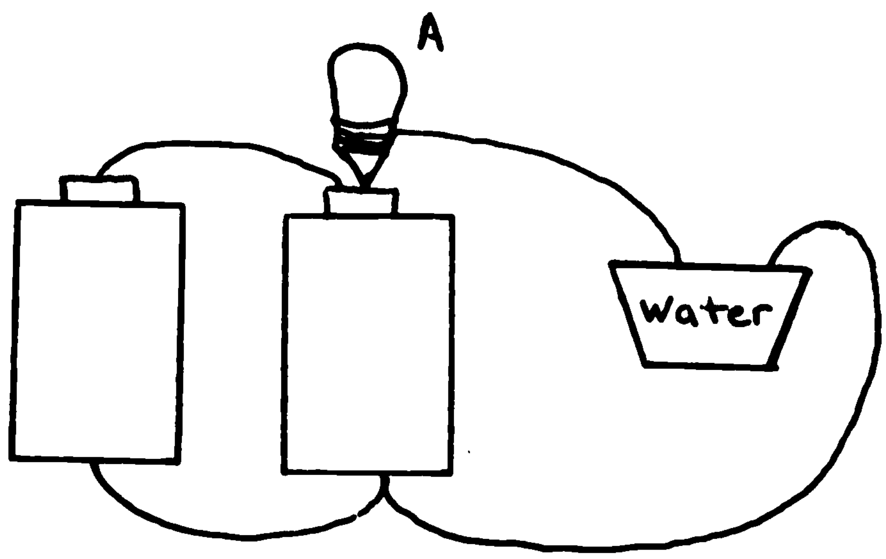
D for Dimmer
N for Not Light

18



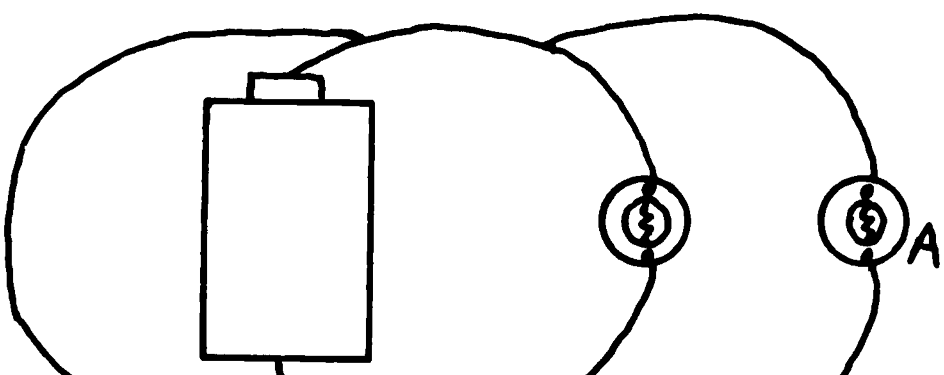
A ______

19



A ______

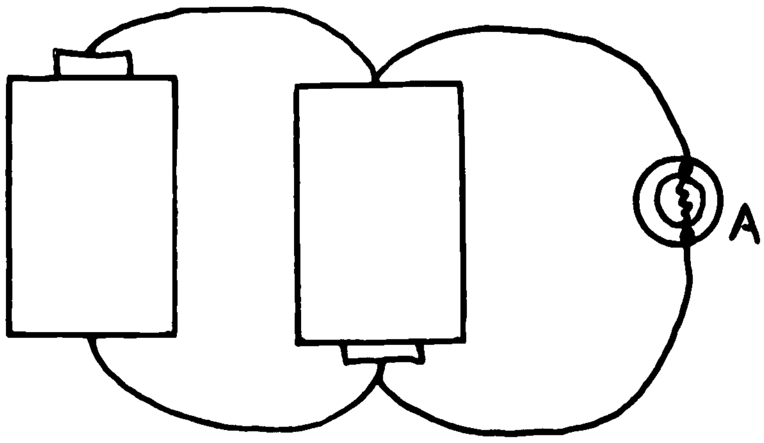
20



A ______

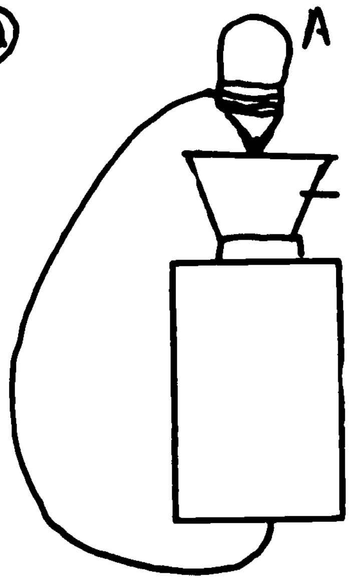
S for Same (or Standard) D for Dimmer
B for Brighter N for Not Light

21)



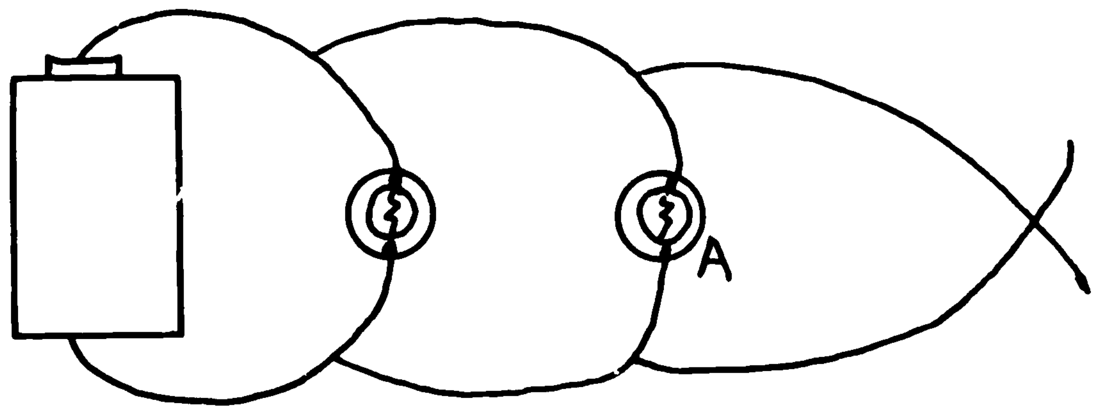
A ______

22)



A ______

23)

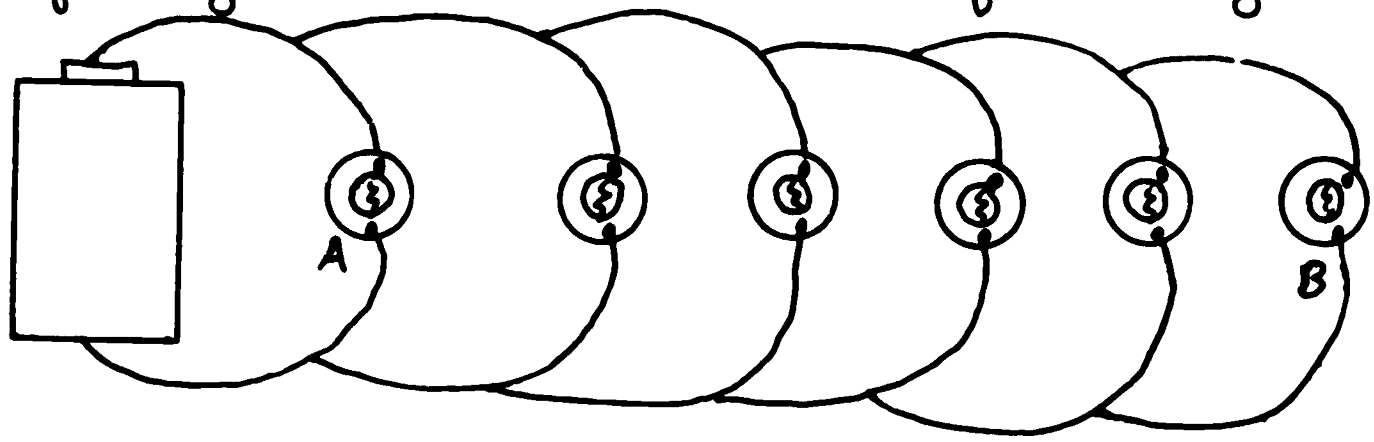


A ______

S for Same (or Standard)
B for Brighter

D for Dimmer
N for Not Light

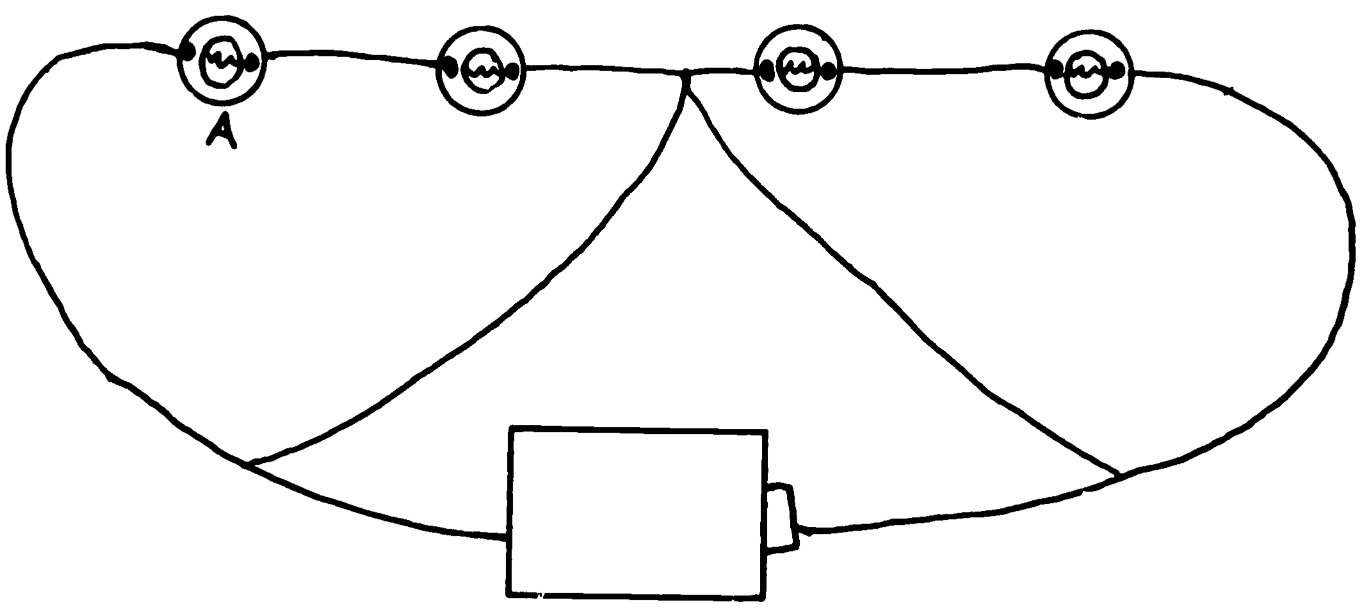
24



A ______

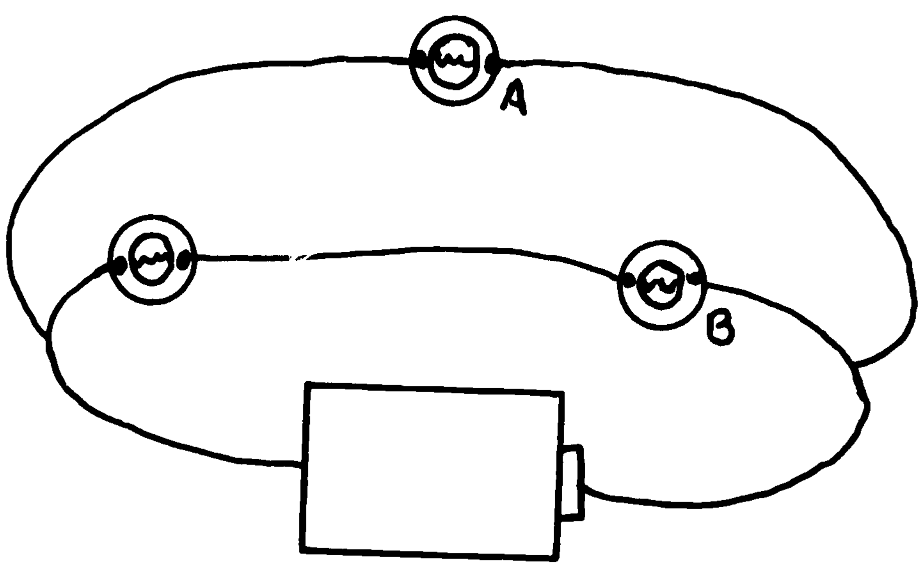
B ______

25



A ______

26



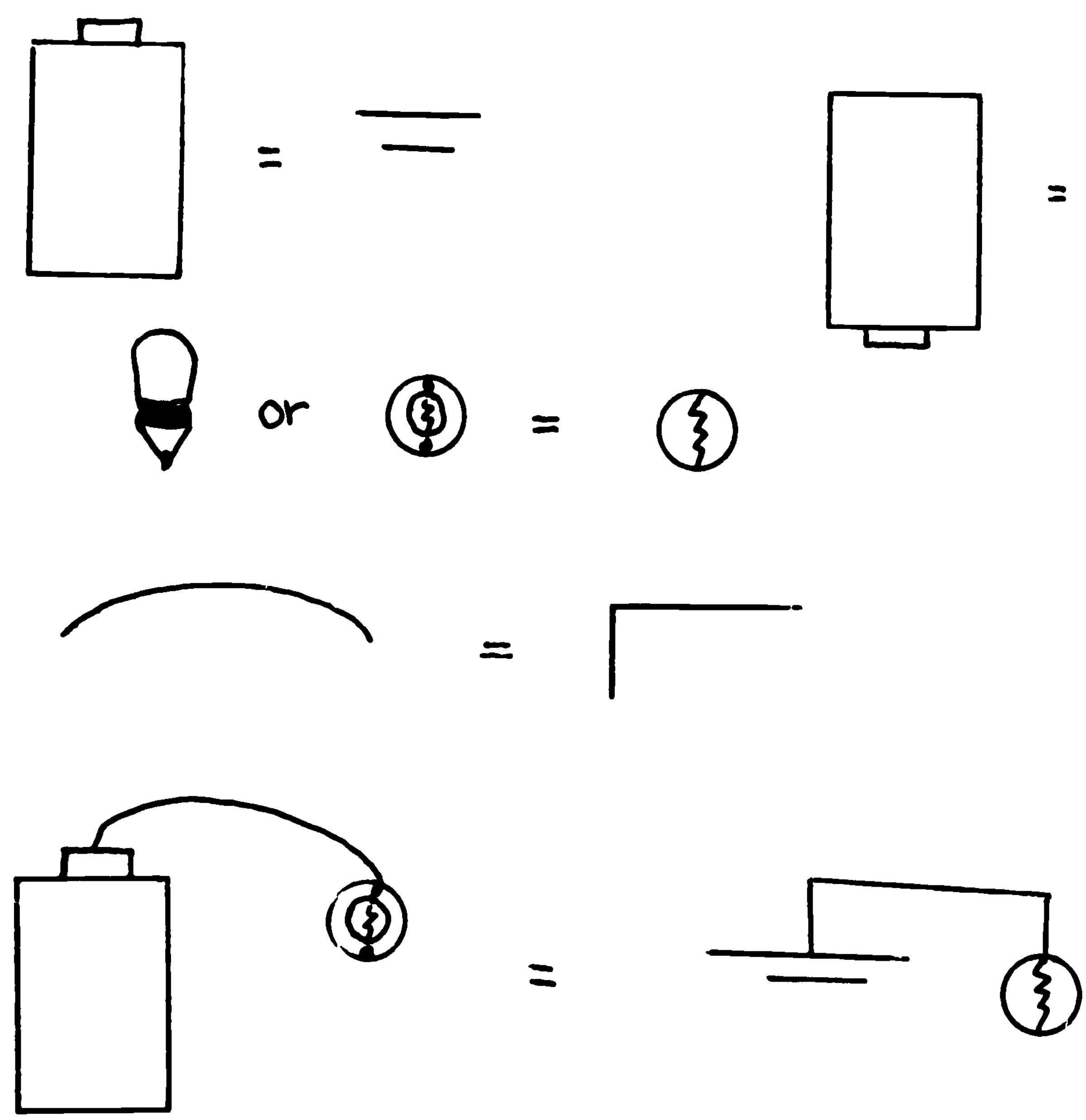
A ______

B ______

END OF PART II:

STOP

New Signs for Part III:



For each setup put

- S for Same (or Standard)
- B for Brighter
- D for Dimmer
- N for Not Light

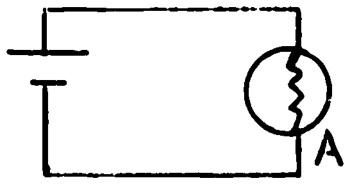
in the blank

PART III

S for Same (or Standard)
B for Brighter

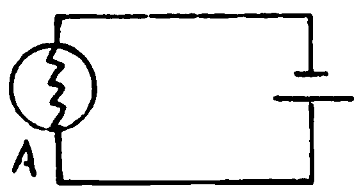
D for Dimmer
N for Not Light

1



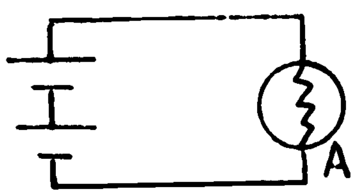
A ______

2.



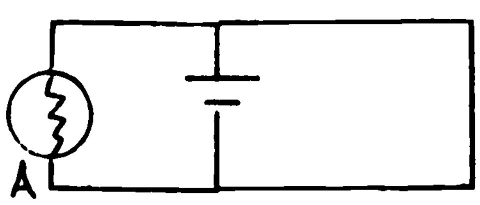
A ______

3



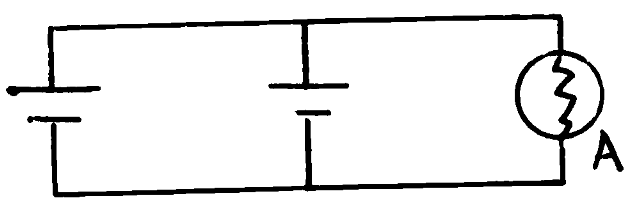
A ______

4.



A ______

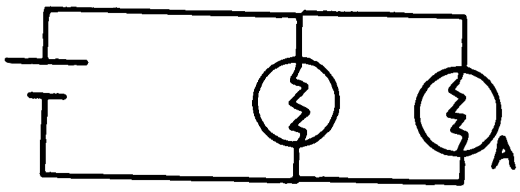
5



A ______

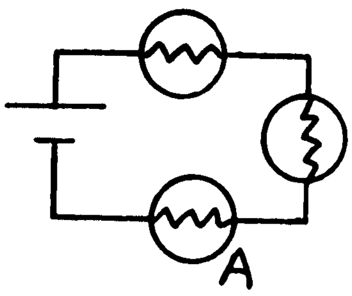
S for Same (or Standard) D for Dimmer
B for Brighter N for Not Light

6.



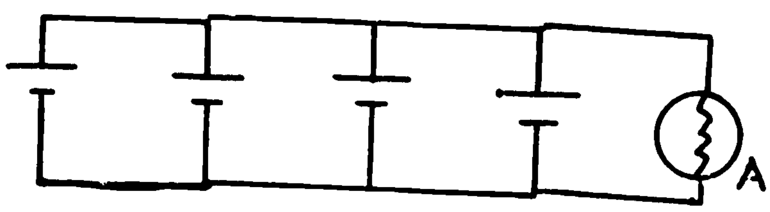
A ______

7.



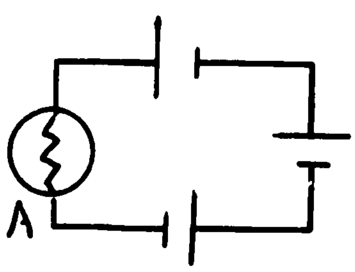
A ______

8.



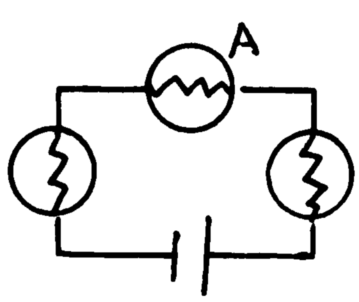
A ______

9.



A ______

10.

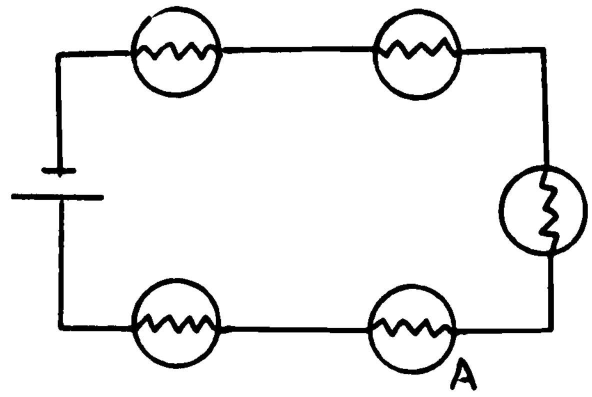


A ______

S for Same (or Standard)
B for Brighter

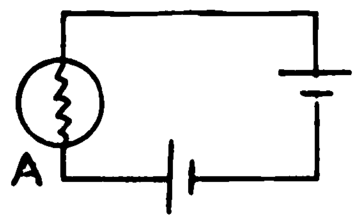
D for Dimmer
N for Not Light

11.



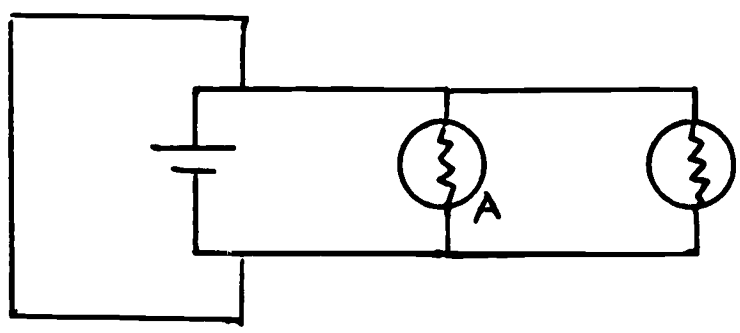
A ________

12.



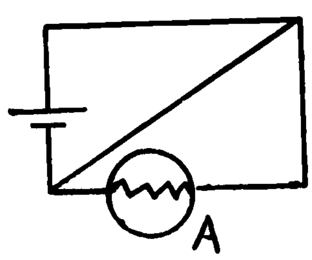
A ________

13.



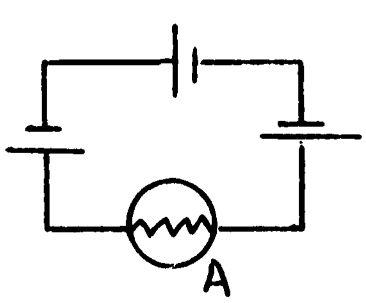
A ________

14.



A ________

15.

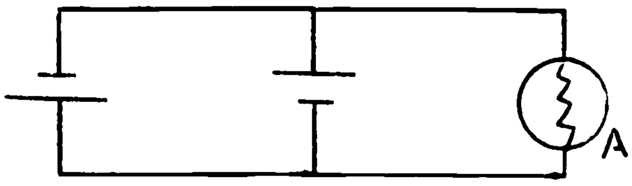


A ________

S for Same or Standard
B for Brighter

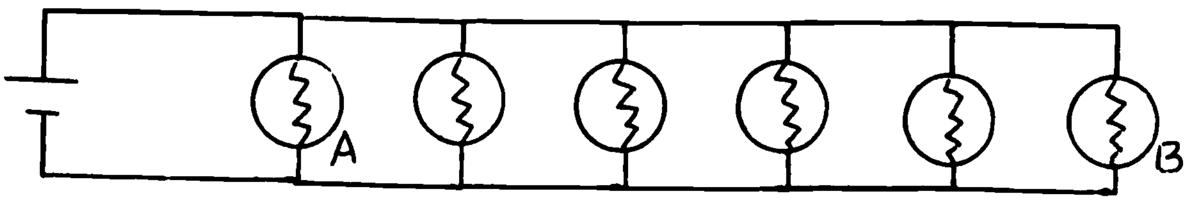
D for Dimmer
N for Not Light

16.



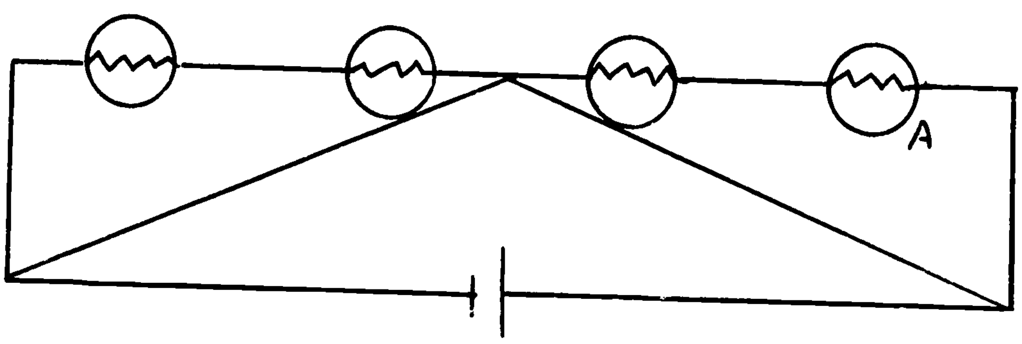
A ______

17.



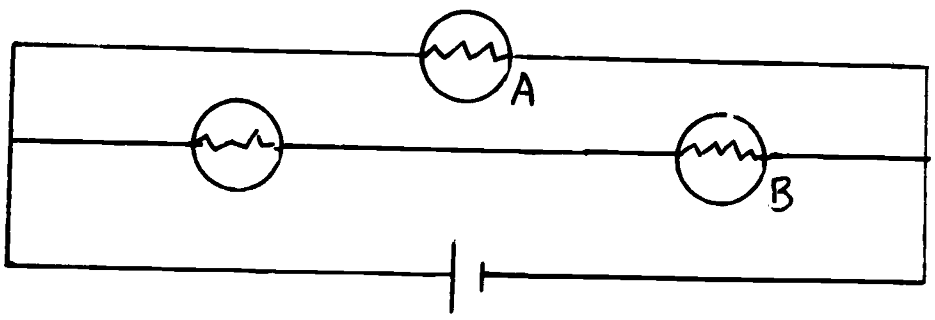
A ______
B ______

18.



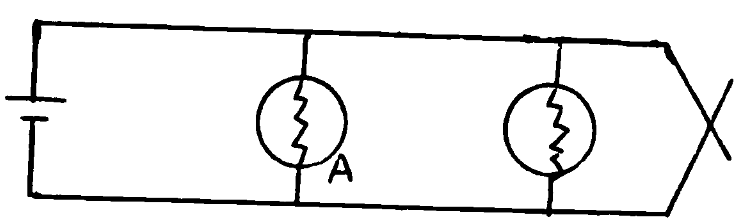
A ______

19.



A ______
B ______

20.



A ______

END OF PART III

STOP

APPENDIX C.

EFFECTS OF A TRAINING PROGRAM ON THE PRODUCTIVE THINKING OF EDUCABLE MENTAL RETARDATES: A FAILURE TO REPLICATE

EFFECTS OF A TRAINING PROGRAM ON THE PRODUCTIVE THINKING OF EDUCABLE MENTAL RETARDATES: A FAILURE TO REPLICATE[1,2]

Milton Budoff, Jean D. Meskin, and Deborah J. Kemler

Cambridge Mental Health Center
Harvard Medical School

Rouse (1965) sought to enhance the productive thinking of educable mentally retarded (EMR) children through exposure to a special curriculum unit aimed at increasing "ideational fluency" through the process of "brainstorming", developing an understanding of principles used in making changes (e.g., size, shape, addition), encouraging originality and improvisation, and providing broader experience with the senses. She developed a thirty lesson unit, and demonstrated positive results from the intervention. The children who received the experimental unit scored markedly higher on a posttest administration of selected subtests of the Minnesota Tests of Creative Thinking (Torrance, 1962) than matched control classes. Rouse (1965) concluded "that a systematic program for the development of productive thinking would enhance significantly both verbal and nonverbal...scores of EMR children..." (p. 669).

The present study sought to replicate Rouse's study. Her findings, if supported, would indicate a strategy of instruction by which verbally restricted children might be encouraged and perhaps taught to think and communicate more

---

1 This research was supported by Grant #32-31-0000-6019 from the Division of Handicapped Children and Youth, U.S. Office of Education.

2 This study could not have been performed without the cooperation of Dr. Vincent P. Conners, Director, Department of Special Classes, the late Mr. Charles P. Ruddy, Associate Superintendent, Mr. George Sawyer, Principal, Miss Frances Mullen and other teachers of the Patrick Gavin Junior High School, Boston, Massachusetts.

freely and more appropriately to the assigned task. Training of this type has been shown to benefit other, much less restricted populations such as college students, and high level personnel in industry (see Rouse, 1965). The successful demonstration of its applicability with these children would be particularly important since markedly improved verbal skills might change significantly the academic status of some of these very low achieving children.

Method

The experimental intervention consisted of the thirty lesson plan units in productive thinking employed by Rouse and described in her manual (1963). The training was administered by a teacher within the school system trained by the second author who was familiar with the materials and who at all times attempted to keep the teacher's instruction consistent with that prescribed by Rouse. The training program continued over a period of six weeks on thirty consecutive school days.

Subjects

Ss were twenty-six students enrolled in the special classes for EMRs in a junior high school in Boston, Massachusetts. Table 1 shows the means and standard deviations for age and IQ for the experimental and control groups.

INSERT TABLE 1 ABOUT HERE

Table 1. Mean CAs and IQs of Experimental and Control Groups in the Replication

GROUPS	CA		IQ	
	$\bar{X}$	SD	$\bar{X}$	SD
EXPERIMENTAL	14-4 mos.	13.2 mos.	69.3	5.7
CONTROL	14-4 mos.	11.6 mos.	71.8	7.7
t - test value	<1		<1	

Of the total of fifteen Ss who received the training intervention, thirteen are included in the final sample, two having been eliminated because of excessive absence. The control sample consisted of thirteen EMRs, matched with the experimental group on pretest scores on the Torrance subtests.

Procedure

As in the Rouse study, two subtests of the Minnesota Tests of Creative Thinking the Product Improvement and the Circles Tasks, were administered individually in one session within a week prior to, and following, the conclusion of the curriculum unit. In the Product Improvement Test, Ss are required to produce suggestions for making a toy dog "more fun to play with." A modification of these instructions was necessary because of the age and maturity of the Ss. Instead of having E ask S how to make the toy dog "more fun to play with," Ss were required to think of ways to make the dog "more fun for an eight year old boy or girl to play with." The Circles Task consisted of two pages containing thirty-six circles on which Ss were required to draw pictures of objects by adding lines both internal and external to the prepared outlines, and adding colors if they wished. Administration of the Circles Task followed the prescribed instructions. The tests were conducted orally and individually. Both the Product Improvement and the Circles Tasks are scored along four dimensions: a "fluency" score represents the quantity of total output, a "flexibility" score indicates the number of different categories of responses, an "originality" score rates the number of atypical responses according to a standardized set of norms; and an "elaboration" score indicates the amount of detail or embellishment in S's responses.

A pretest-posttest design was employed to evaluate the effects of the unit. The same forms of both tests were used in the two administrations. All tests were scored by Torrance's staff at the University of Georgia. The scorers were not aware of the experimental status of the Ss.

In addition to this test data, the teacher also made records of the extent of participation of the experimental Ss in the training session. For each day of instruction, each S was rated on the number of contributions he made to the class discussion.

Results

Means, ranges, and standard deviations for pretest, posttest, and gain scores of the experimental and control groups are presented in Tables 2 and 3 for the Product Improvement and Circles Tasks respectively. The results of the t-tests for differences in gain are also indicated on these tables.

INSERT TABLES 2 & 3 ABOUT HERE

With the exception of a gain in originality on the Circles Task ($p<.01$), no other differences in gain were significant. Thus, the major results of the present study are not congruent with Rouse's findings. Furthermore, the significant difference in gain in originality on Circles in the present study must be qualified by the observation that the control group appears to have decreased its mean score from

Table 2. Differences between groups on the Product Improvement Task prior to, and following, the class unit.

	Pre-test	Post-test	Gain	Pre-test	Post-test	Gain	Gain t-value
	Exp. Group (N=13)			Control Group (N=13)			
FLUENCY							
Mean	11.08	16.31	3.46	11.77	13.85	2.08	
Range	4 to 29	4 to 38	-1 to 11	4 to 20	4 to 29	-16 to 12	
SD	7.85	8.69	3.55	4.56	6.93	7.76	<1
FLEXIBILITY							
Mean	5.00	5.62	.62	5.23	4.85	-.38	
Range	3 to 10	1 to 11	-4 to 4	1 to 10	2 to 7	-5 to 5	
SD	2.32	2.05	2.03	2.12	1.74	2.64	<1
ORIGINALITY							
Mean	12.30	13.00	.70	11.08	12.85	1.77	
Range	1 to 30	1 to 26	-7 to 13	1 to 19	1 to 27	-16 to 16	
SD	8.75	8.49	4.76	4.81	7.68	9.38	<1
ELABORATION							
Mean	1.62	2.15	0.53	2.69	2.92	.23	
Range	0 to 6	0 to 7	-2 to 4	0 to 9	0 to 10	-9 to 10	
SD	1.64	2.29	1.55	2.58	3.30	4.19	<1

Table 3. Differences between groups on the Circles Task prior to, and following, the class unit.

	Pre-test	Post-test	Gain	Pre-test	Post-test	Gain	Gain t-value
	Exp. Group (N-13)			Control Group (N-13)			
FLUENCY							
Mean	7.69	10.62	2.93	8.23	10.69	2.54	
Range	5 to 15	5 to 21	-4 to 10	3 to 15	5 to 20	-4 to 11	
SD	3.17	5.23	4.13	3.81	4.28	3.60	<1
FLEXIBILITY							
Mean	6.38	8.46	2.08	6.38	7.77	1.38	
Range	2 to 12	1 to 16	-4 to 7	3 to 12	3 to 12	-3 to 7	
SD	3.01	4.43	3.26	2.70	2.94	2.76	<1
ORIGINALITY							
Mean	11.46	15.62	4.16	12.69	8.46	-5.00	
Range	1 to 40	0 to 42	-28 to 32	0 to 31	0 to 21	-27 to 5	1.98*
SD	11.35	14.70	14.20	8.83	5.50	5.73	
ELABORATION							
Mean	17.62	22.00	2.38	19.54	22.85	3.31	
Range	9 to 26	16 to 34	-8 to 14	9 to 40	9 to 39	-14 to 23	<1
SD	5.12	4.69	6.79	8.81	9.18	10.83	

* $p < .025$

pretest to posttest to the same extent that the scores of the experimental group improved. Thus, the significance may be due to the loss in the control group at least as much as it is attributable to increased ability in the experimental group. Although the finding of a considerable reduction in score from pre-to posttest is surprising since one might postulate some practice effect operating in the second administration, Torrance (1966) states that scores are equally likely to rise or fall from test to retest.

The use of multiple t-tests to analyze differences within a given set of data is a problematical procedure from a statistical point of view since the probability of a Type 1 error is increased. These were run in order to maintain the parallel between the present study and that of Rouse. In addition, a discriminant analysis was conducted both on posttest and on gain scores. The analysis was run with two groups (experimental and control) and eight variables (four dimensions on the two tasks). Separate analyses for posttest and gain scores revealed no significant $\underline{F}$ ratios on the individual variables.

Correlation coefficients between scores on the two subtests (Product Improvement and Circles) were run independently for the four dimensions in order to determine the degree of relationship between the two tasks. While the scores for fluency and flexibility (r=.33) are related, the scores for originality (r=.14) and for elaboration (r=.10) on the two tasks are not significantly related. The data employed in these analyses were the pretest scores of a total of fifty-four subjects, the total sample to which the tests were administered prior to the construction of experimental and control groups.

On the assumption that extent of participation in the classroom training session should be closely associated with gain scores on the tests, mean measures of participation for each experimental S were correlated with pretest, posttest, and gain scores for all dimensions on the two tasks. The results are shown in Table 4, and, in general, do not support the hypothesis.

INSERT TABLE 4 ABOUT HERE

Although the results as a whole are difficult to interpret, the negative correlations between extent of participation and measures of gain in fluency and flexibility on the Circles Task are striking in that they are in a direction directly opposite to the expected relationship.

The data from this replication do not lend support to the hypothesis that training in productive thinking on Rouse's unit markedly improved the ability of EMRs to perform on selected subtests of the Minnesota Tests of Creative Thinking.

Discussion

The discrepancies between the Rouse's findings and the present study are probably attributable to differences in the subject populations rather than to any differences in the intervention procedure. As far as can be determined, the classroom atmosphere in the present study was as conducive to "productive thinking" as that achieved by Rouse's instructors.

Table 4. Correlations with pre-test, gain, and post-test scores on the two creative thinking tasks (N=14)

	Pre-test	Gain	Post-test
PRODUCT IMPROVEMENT			
Fluency	.30	.53*	.55*
Flexibility	-.10	.43	.28
Originality	.36	-.05	.35
Elaboration	.62**	-.44	.09
CIRCLES TASK			
Fluency	.28	-.33	-.13
Flexibility	.16	-.46	-.24
Originality	-.49*	.22	-.17
Elaboration	.08	-.08	-.04

* $p<.05$
** $p<.01$

The first indication that the Ss in the present study differed from those in Rouse's sample is provided by a comparison of the pretest scores of the two groups. Mean pretest scores for the two samples are shown in Table 5.

INSERT TABLE 5 ABOUT HERE

It may be seen that prior to the intervention, the Boston EMRs were superior performers to their South Carolina counterparts on the two subtests of the Minnesota Tests of Creative Thinking. At first glance, such performance differences seem to be attributable to the age discrepancy of the two experimental samples. The mean CA of the present sample was higher than that of Rouse's sample, and the SD was smaller. One might argue, then, either that Rouse's unit was not a suitable intervention for the present sample because of their greater CA, or that those subtests selected to test for gain were not appropriate instruments for use with this older non-middle class sample.

The former argument that the CA differences alone were important in determining susceptibility to the unit may also be discounted. With Rouse's cooperation, it was possible to match a group of her experimental Ss with the present sample on the basis of CA and IQ. A comparison of the mean pretest, posttest, and gain scores of the two groups on the Minnesota tests are shown in Table 6.

Table 5. Comparison of EMRs in this replication and Rouse's study.

	PRESENT SAMPLE	ROUSE'S SAMPLE	t
PRODUCT IMPROVEMENT			
Fluency	11.08	4.00	2.68*
Flexibility	5.00	2.15	3.52**
Originality	12.30	4.31	2.63*
Elaboration	1.62	1.15	0.66
CIRCLES TASK			
Fluency	7.69	5.92	1.23
Flexibility	6.38	4.15	1.90
Originality	11.46	3.85	2.20*
Elaboration	17.62	10.46	2.26*

* $p < .05$
** $p < .01$

INSERT TABLE 6 ABOUT HERE

Consonant with the results of the two studies, Rouse's sample gained considerably more than the Boston Ss. However, a comparison of posttest scores shows the two groups to be functioning at the same level following the intervention. This finding is due to the large difference in performance of the two groups on the pretest, with the Boston Ss superior to Rouse Ss on all measures, significantly so on six of the eight measures. Thus, it does not appear to be the age difference as such which accounts for the discrepant results of the two studies. Equivalent CA and IQ samples from the two studies performed dissimilarly on the Minnesota tests before any experimental intervention. The differences between the two experimental groups appear to be more complex.

Although the Ss in the present study were drawn from the population of "culturally deprived", it should be noted that they have been given considerable attention during the last two years by the present group of investigators. Over this period, all Ss have received several types of personality and learning tests, and all Ss have had considerable contact with numerous investigators. Rouse reported, in personal communication, that pretesting and posttesting were her only contacts with her Ss. Both samples are from similar social backgrounds. A "test-wise" attitude may account for the initially higher scores of the Boston Ss on the Product Improvement and Circles Tasks. It may be because they were so

Table 6. A comparison of the scores achieved by a matched sub-group of Rouse's experimental sample and the sample in the present study.

	Present Sample (N=13)			Rouse's Sample (N=13)		
	Pre-test	Post-test	Gain	Pre-test	Post-test	Gain
PRODUCT IMPROVEMENT						
Fluency	11.08	16.31	3.46	4.00	11.69	7.69
Flexibility	5.00	5.62	0.62	2.15	5.54	3.39
Originality	12.30	13.00	0.70	4.31	9.69	5.38
Elaboration	1.62	2.15	0.53	1.15	3.69	2.54
CIRCLES TASK						
Fluency	7.69	10.62	2.93	5.92	9.85	3.93
Flexibility	6.38	8.46	2.08	4.15	7.69	3.54
Originality	11.46	15.62	4.16	3.85	9.23	5.38
Elaboration	17.62	22.00	2.38	10.46	18.77	8.31

familiar with an evaluation process they had come to see as nonthreatening. By contrast, EMR populations who do not have this prior testing exposure tend to be suspicious and less responsive. The higher posttest scores of the South Carolina sample may reflect, in part, the lessened suspicion towards the previously unfamiliar examiner. Also, there may be a ceiling effect on these selected subtests of the Minnesota Tests of Creative Thinking. It has been observed frequently that non-middle class children are not motivated to give large numbers of responses in the same striving manner characteristic of middle class children. They do not seem to be intrigued by the game without clearly specified tangible reinforcement (Terrell, Durkin & Wiesley, 1959; Zigler & Delabry, 1962). This ceiling effect may explain the relatively smaller increment in score following presentation of the unit to the Boston School System

It also should be mentioned that the persons working with the lessons composing the unit did feel strongly that the tasks presented were too young for these adolescent aged EMRs, though they may be appropriate for their mental age peers. That is, these adolescent EMRs' interest level and knowledge more closely approximated their CA peers' interests and knowledge. They found it difficult to become involved with problems that might engage younger children. It may be wise to be much less literal about assuming that educational materials for EMRs should be directed to their MA level.

It appears to the present authors that any attempt to generalize the results of the Rouse study must be conservative and cautious. This conclusion is based both on a consideration of the present failure to replicate her empirical findings

and also on a critical rethinking of the Rouse study itself. There is a danger that overzealous readers impressed with Rouse's success in increasing scores on two tests purporting to measure "creativity" will not perceive the important differences between augmenting scores on a test and improving the underlying ability which the test is supposed to be measuring. Rouse succeeded in doing the former but there is as yet no evidence that she accomplished the latter. Rouse herself is cognizant of the difference as she emphasizes the importance of further studies to evaluate the degree to which the training intervention generalizes to areas not covered by the unit. As of now, no such studies have been conducted.

REFERENCES

Rouse, Sue T. Effects of a training program on the productive thinking of educable mental retardates. Unpublished doctoral dissertation, George Peabody College, 1963.

Rouse, Sue T. Effects of a training program on the productive thinking of educable mental retardates. American Journal of Mental Deficiency, 1965, 69 (5), 666-673

Terrell, G., Jr., Durkin, Kathryn, & Wiesley, M. Social class and the nature of the incentive in discrimination learning. Journal of Abnormal and Social Psychology, 1959, 59, 270-272.

Torrance, E. P. Guiding creative talent. Englewood Cliffs, N. J.: Prentice-Hall, 1962(a).

Torrance, E. P. Torrance tests of creative thinking: Directions manual and scoring guide. Princeton, New Jersey: Personnel Press, Inc., 1966.

Zigler, E. & De Labry, J. Concept switching in middle class, lower class and retarded children. Journal of Abnormal and Social Psychology, 1962, 65, 267-273.

APPENDIX D.

RATING SCALES ON STUDENTS' CLASSROOM
BEHAVIOR FOR OBSERVERS AND CLASSROOM TEACHERS

Rating for Classroom Observers*

Please rate each child on the following variables:

Motivation

1. Tries out new ideas - excited about findings, not necessarily all period - even it it's just a couple of times during lesson.

2. Work holds his attention.

3. Has to be pushed into doing work - it may hold his attention once started.

Attention

1. Works pretty steadily through period - can spend intervals in non-work but on whole is working.

2. Fifty percent of time devoted to work - may start in middle of period or alternate periods of work with non-work.

3. Works at task for short period of time (10 - 15 minutes) or not at all - Starts in beginning and stops - starts at end of period.

Need to be Acknowledged

1. Runs up to teacher with work - shows friends - runs to board - raises hand - or negatively - yelling out - disrupting class.

2. Shows friend if they come to his desk - or once or twice during period will show teacher work - wanders around room sometimes.

3. Works by self or sits quietly by self.

Role in Class

A-C - Active-Constructive - (goes to board, helps other students, helps teacher, volunteers answers).

P-C - Passive Constructive - works quietly at desk, listens to teacher explaining things to other students, or to class as a whole.

Average - no significant contribution to the class in either direction.

P-D - Passive disruptive - disruptive in a quiet way, or encourages disruptive child by smiling at him, watching or assisting him.

A-D - Active-Disruptive - clowns around, yells out in class, interrupts other children who are working.

*Used only in main study 1966-67.

Rating for Teacher - Understanding Electricity*

Please rate each child 1-5 according to which category best fits him.

I. Ability

1. Excellent ability in class - easily grasps new materials and retains knowledge well.

2. Good ability in learning and retaining knowledge.

3. Average ability and has a little difficulty with work.

4. Below average ability has some difficulty managing classwork.

5. Very poor ability -- barely manages or is not able to manage demands of class.

II. Cooperation

1. Very cooperative - volunteers to assist, helps other students, immediately follows through on your requests for them to do board work, written material, etc.

2. Good cooperation - does not volunteer but does meet your demands immediately.

3. Average cooperation -- meets your demands though not necessarily immediately or willingly.

4. Below average cooperation -- some difficulty in meeting requirements of the classroom -- sometimes disruptive.

5. Very uncooperative -- does not, or balks at meeting requirements of classroom, noisy or disruptive as a rule.

III. Participation in Class Discussion

1. Very frequent

2. Frequent

3. Average

4. Below average

5. Very far below average

*Lecture-Demonstration unit used in conducting study 1967-68.

IV. Interest and Enthusiasm

1. Participates with great enthusiasm, interest and curiosity.

2. Is generally enthusiastic and curious.

3. Interested and enthusiastic from time to time - vacillates.

4. Rarely shows interest and enthusiasm towards subject matter or experiments.

5. Indifferent and bored - no enthusiasm evident.

Ratings for Teacher - Experiments in Electricity*

Please rate each child 1-5 according to which category best fits him.

I. Ability

1. Excellent ability in class - easily grasps new materials and retains knowledge well.
2. Good ability in learning and retaining knowledge.
3. Average ability and has a little difficulty with work.
4. Below average ability has some difficulty managing classwork.
5. Very poor ability -- barely manages or is not able to manage demands of class.

II. Productivity and Application to Work

1. Sees problem through to end and goes on to extra activity.
2. Completes task successfully though does not go on to exta work.
3. Performs task satisfactorily though sporadically.
4. Has trouble completing task though does work on it.
5. Has a great deal of trouble, perhaps does not make a start on it.

III. Cooperation

1. Very cooperative - volunteers to assist, helps other students, immediately follows through on your requests for them to do board work, written material, etc.
2. Good cooperation - does not volunteer but does meet your demands immediately.
3. Average cooperation -- meets your demands though not necessarily immediately or willingly.
4. Below average cooperation -- some difficulty in meeting requirements of the classroom -- sometimes disruptive.
5. Very uncooperative -- does not, or balks at meeting requirements of classroom, noisy or disruptive as a rule.

*Manipulation unit used in main and concluding studies, 1966-7,1967-8. Title only used in concluding study.

APPENDIX E

ELECTRICITY POSTTEST SCORES FOR MAIN STUDY (1966-7) COVARIED BY IQ

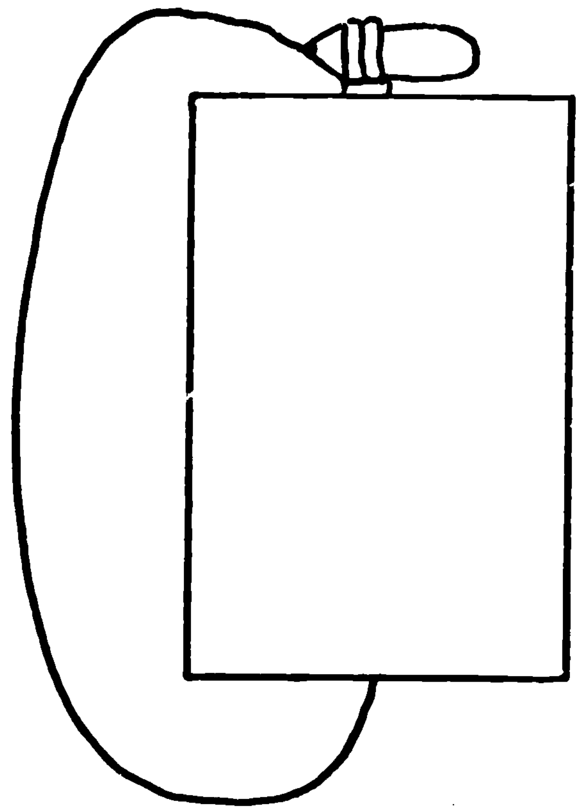
The results of these analyses are summarized in Notes 1 and 3 in Section IV, pp. 25-26 and 28-29, respectively.

APPENDIX F

THE VERBAL REASONS SECTION OF THE EVALUATION INSTRUMENT

PINK BULB

THICK COPPER WIRE



IN THIS SET UP:

1. THE ELECTRIC POWER CAN'T TRAVEL BECAUSE THE BULB IS ON ITS SIDE.

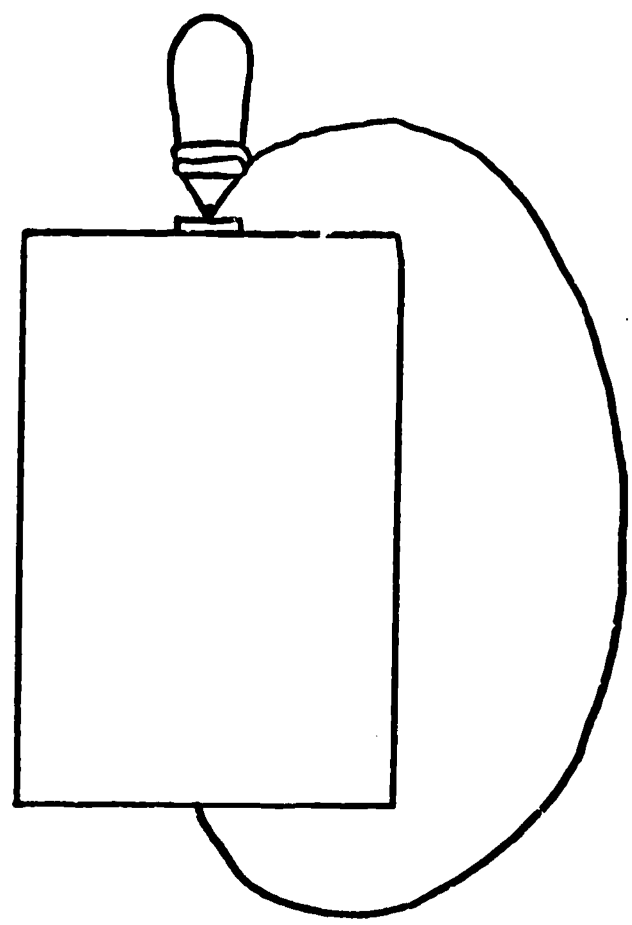
2. THE ELECTRIC POWER HAS TROUBLE TRAVELING BECAUSE YOU NEED MORE THAN ONE BATTERY TO LIGHT THE BULB.

3. THE ELECTRIC POWER TRAVELS EASILY BECAUSE ALL THE RIGHT PLACES ARE CONNECTED.

4. THE ELECTRIC POWER TRAVELS EASILY BECAUSE THE BATTERY IS RIGHT SIDE UP.

PINK BULB

THICK NICHROME WIRE



IN THIS SET UP:

1. THE ELECTRIC POWER TRAVELS EASILY BECAUSE THERE IS NO BULB-HOLDER.

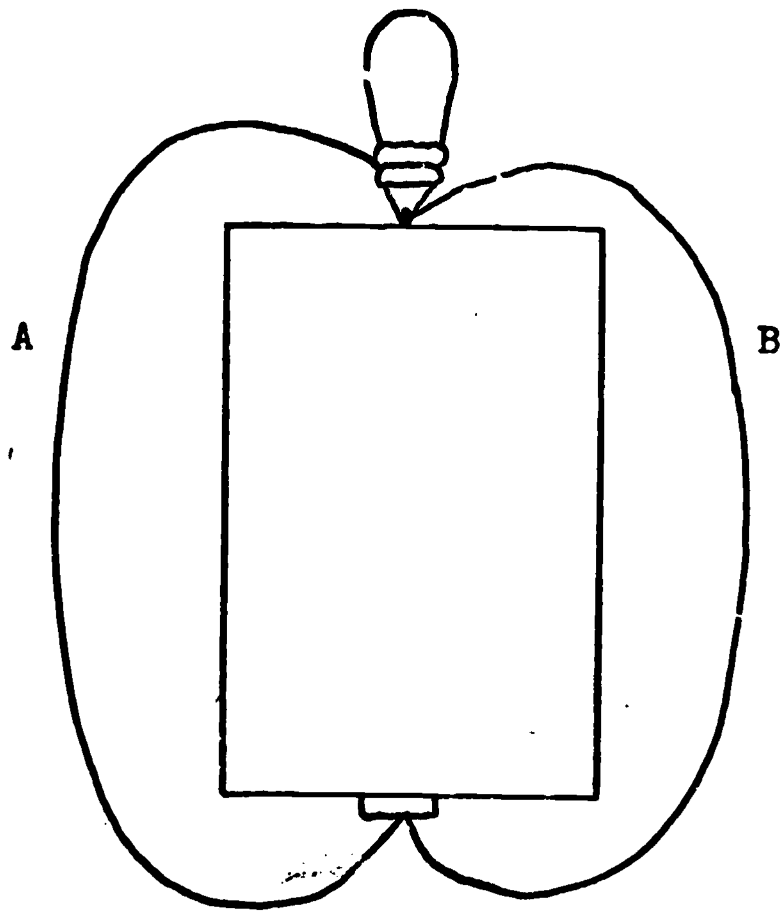
2. THE ELECTRIC POWER HAS TROUBLE TRAVELING THROUGH THE WIRE.

3. THE ELECTRIC POWER CAN'T PASS THROUGH THE WIRE AT ALL.

4. THE ELECTRIC POWER TRAVELS EASILY THROUGH THE WIRE.

PINK BULB

THICK COPPER WIRE



IN THIS SET UP:

1. THE ELECTRIC POWER TRAVELS BOTH EASY PATHS.

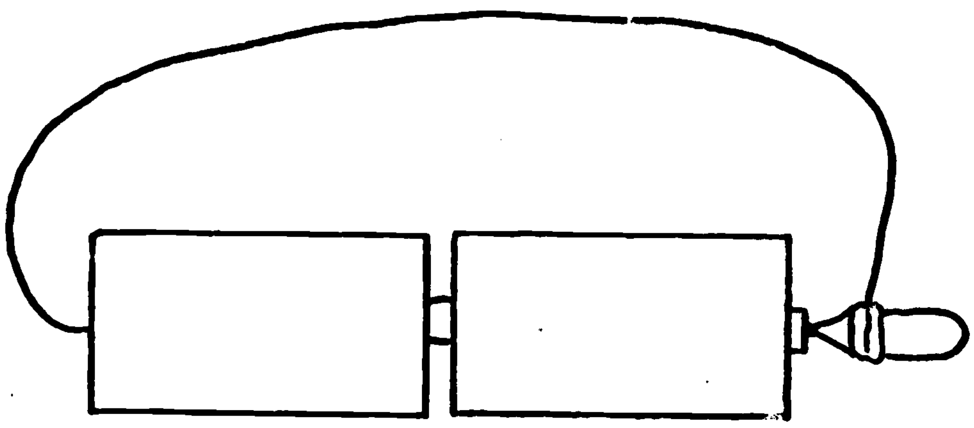
2. THE ELECTRIC POWER TAKES THE EASIER PATH A.

3. THE ELECTRIC POWER TAKES THE EASIER PATH B.

4. THE ELECTRIC POWER CANNOT TRAVEL EITHER PATH BECAUSE THE BATTERY IS UPSIDE DOWN.

pink bulb

THICK COPPER WIRE



IN THIS SET UP:

1. EACH BATTERY GIVES OUT ONLY ONE-HALF ITS ELECTRIC POWER.

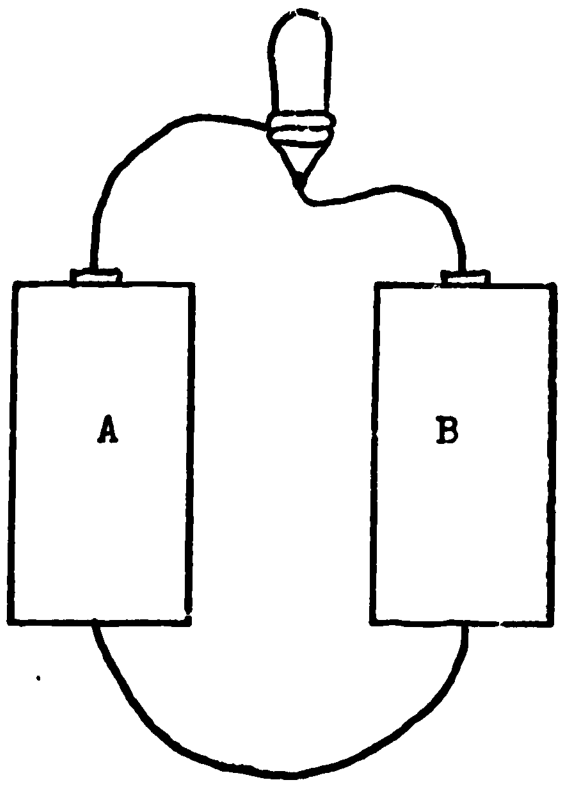
2. THE 2 BATTERIES CANCEL EACH OTHER OUT.

3. NEITHER BATTERY WILL WORK BECAUSE THEY ARE BOTH ON THEIR SIDES.

4. THE 2 BATTERIES ADD THEIR POWER TOGETHER.

PINK BULB

THICK COPPER WIRES



IN THIS SET UP:

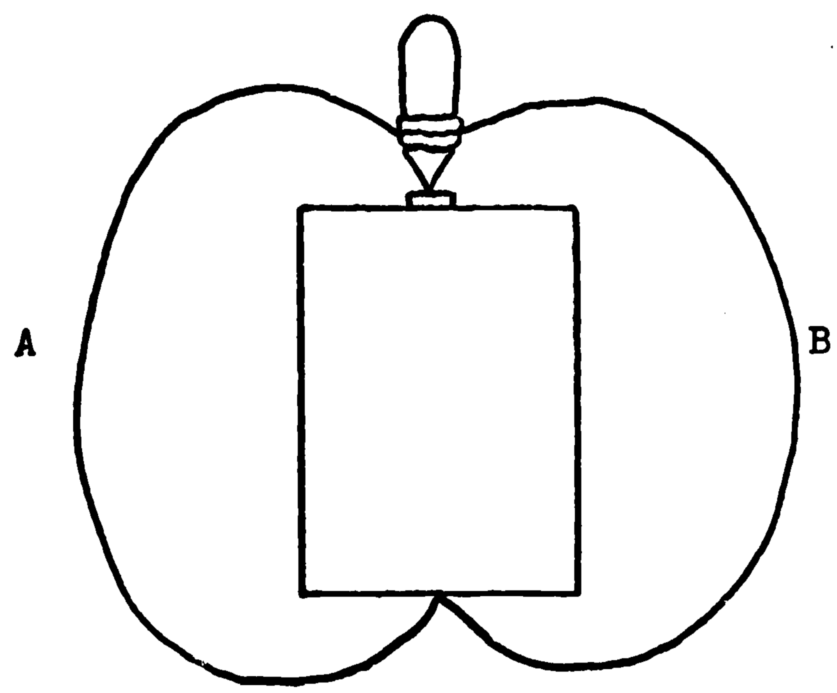
1. THE 2 BATTERIES CANCEL EACH OTHER OUT.

2. BATTERY A IS NOT USED UNTIL BATTERY B RUNS DOWN.

3. THE 2 BATTERIES ADD THEIR POWER TOGETHER.

4. TOO MANY BATTERIES BLOW OUT THE BULB.

THICK COPPER WIRES



IN THIS SET UP:

1. THE ELECTRIC POWER DOESN'T GO THROUGH EITHER PATH BECAUSE IT CAN'T GET THROUGH THE COPPER WIRE.

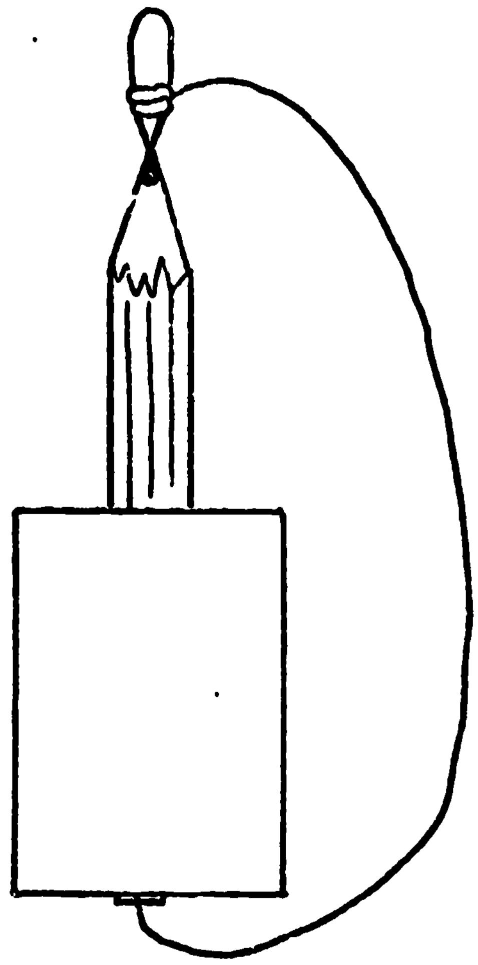
2. THE ELECTRIC POWER TAKES THE EASIER PATH A.

3. THE ELECTRIC POWER TAKES THE EASIER PATH B.

4. THE ELECTRIC POWER TAKES BOTH EASY PATHS.

PINK BULB

THICK COPPER WIRE



IN THIS SET UP:

1. THE ELECTRIC POWER CAN'T PASS THROUGH THE PENCIL.

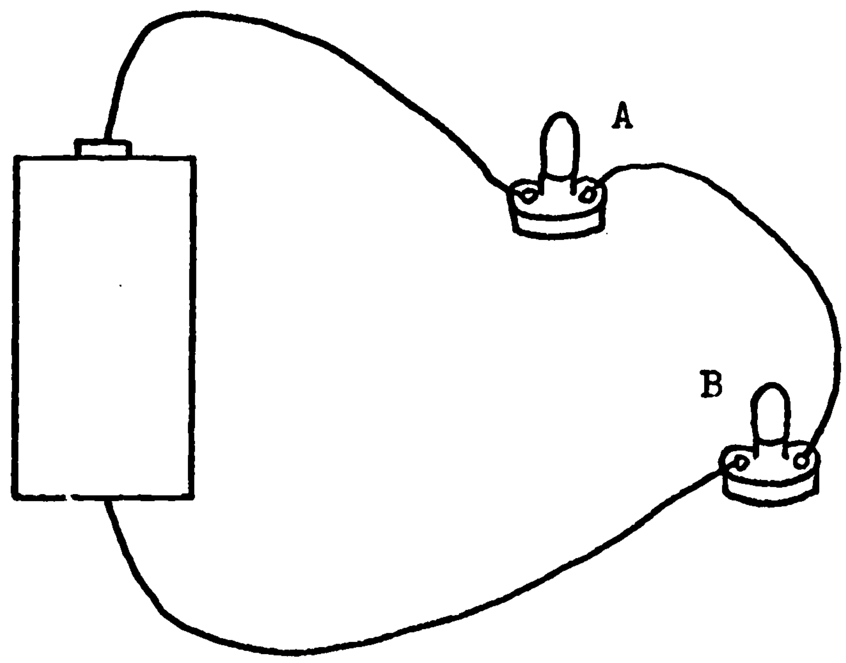
2. THE ELECTRIC POWER PASSES EASILY THROUGH THE PENCIL.

3. THE ELECTRIC POWER CAN'T WORK WITHOUT ANOTHER WIRE.

4. THE ELECTRIC POWER TRAVELS EASILY BECAUSE THE BATTERY IS UPSIDE DOWN.

ALL PINK BULBS

THICK COPPER WIRES



IN THIS SET UP:

1. EACH BULB GETS ELECTRIC POWER AT THE SAME RATE THAT ONE STANDARD BULB DOES.

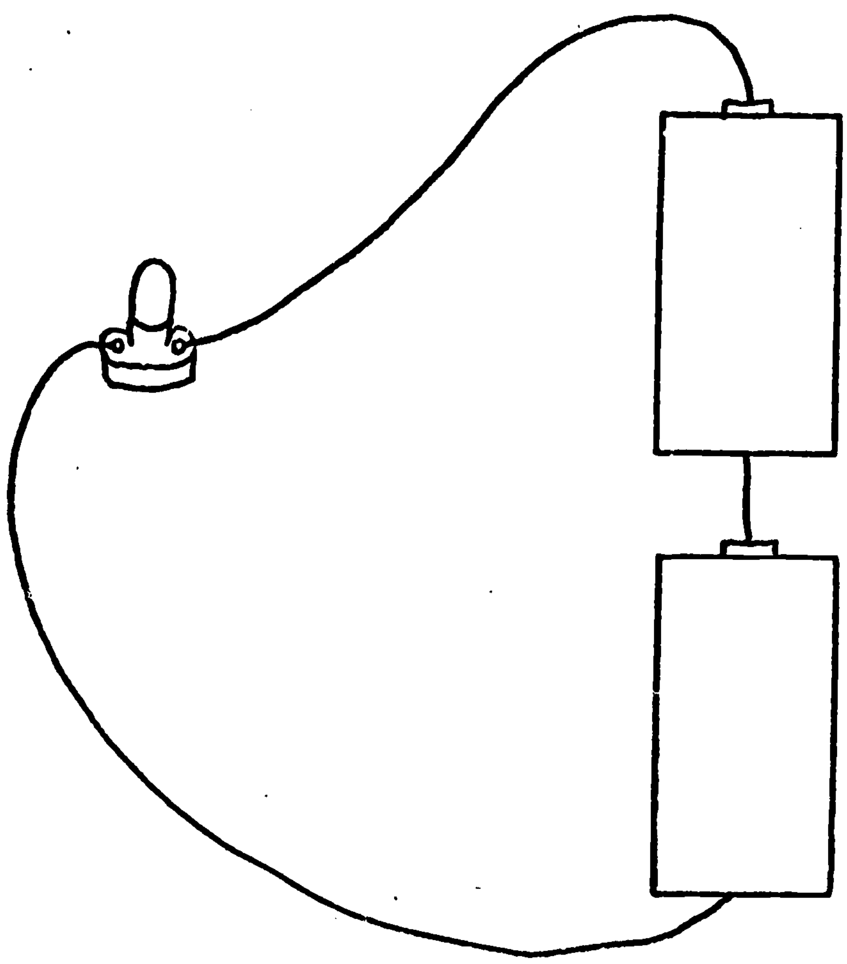
2. EACH BULB GETS ELECTRIC POWER AT A FASTER RATE THAN ONE STANDARD BULB DOES.

3. EACH BULB GETS ELECTRIC POWER AT A SLOWER RATE THAN ONE STANDARD BULB DOES.

4. BULB A GETS POWER FROM THE BATTERY AT A FASTER RATE THAN BULB B.

PINK BULB

THICK COPPER WIRES



IN THIS SET UP:

1. THE WIRE BETWEEN THE BATTERIES SHORT-CIRCUITS THE POWER.

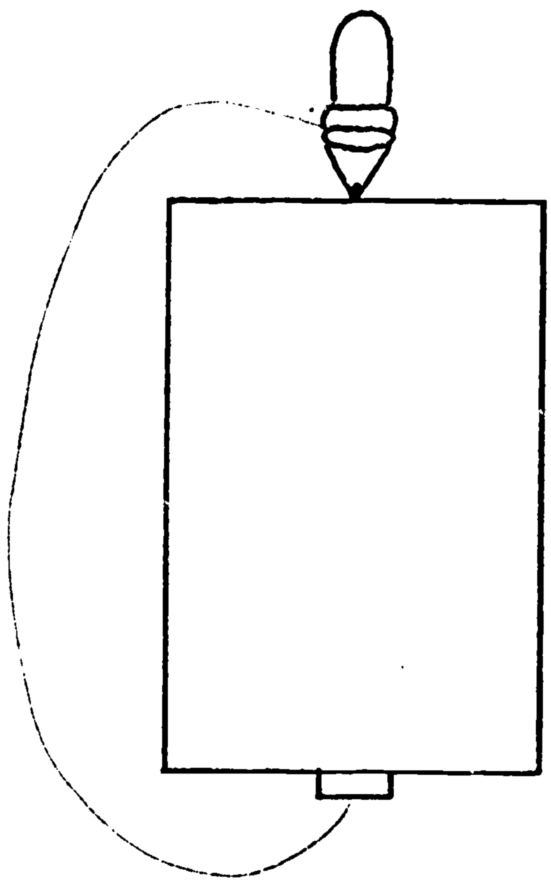
2. EACH BATTERY GIVES OUT ONLY ONE HALF OF ITS ELECTRIC POWER.

3. THE 2 BATTERIES CANCEL EACH OTHER OUT.

4. THE 2 BATTERIES ADD THEIR POWER TOGETHER.

PINK BULB

THIN COPPER WIRE



IN THIS SET UP:

1. THE ELECTRIC POWER HAS A HARD TIME PASSING THROUGH THE WIRE.

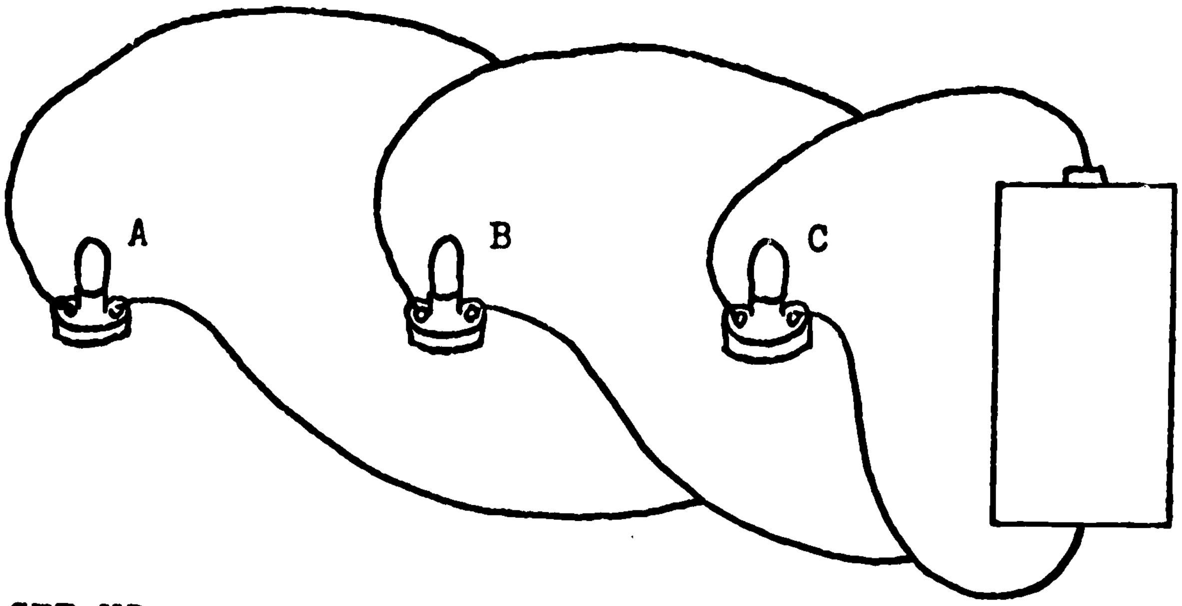
2. THE ELECTRIC POWER CAN'T EVEN GET OUT OF THE BATTERY BECAUSE ITS UPSIDE DOWN.

3. THE ELECTRIC POWER PASSES EASILY THROUGH THE WIRE.

4. THE ELECTRIC POWER CAN'T PASS AT ALL THROUGH THE WIRE.

ALL PINK BULBS

THICK COPPER WIRES



IN THIS SET UP:

1. EACH BULB GETS POWER AT A SLOWER RATE THAN ONE STANDARD BULB DOES.

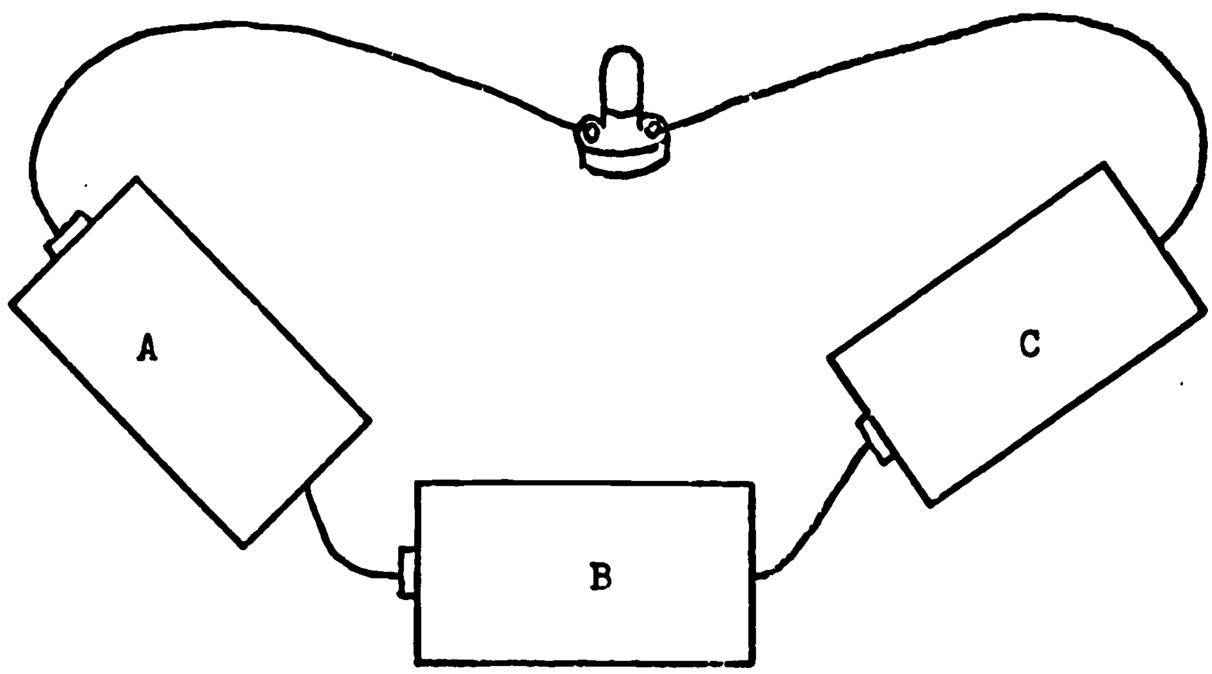
2. EACH BULB GETS POWER AT THE SAME RATE THAT ONE STANDARD BULB DOES.

3. BULB A DOESN'T GET ANY POWER BECAUSE IT IS TOO FAR AWAY FROM THE BATTERY.

4. BULB C GETS POWER AT A FASTER RATE THAN BULBS A AND B DO.

PINK BULB

COPPER WIRES



IN THIS SET UP:

1. BATTERIES A AND B CANCEL EACH OTHER OUT.

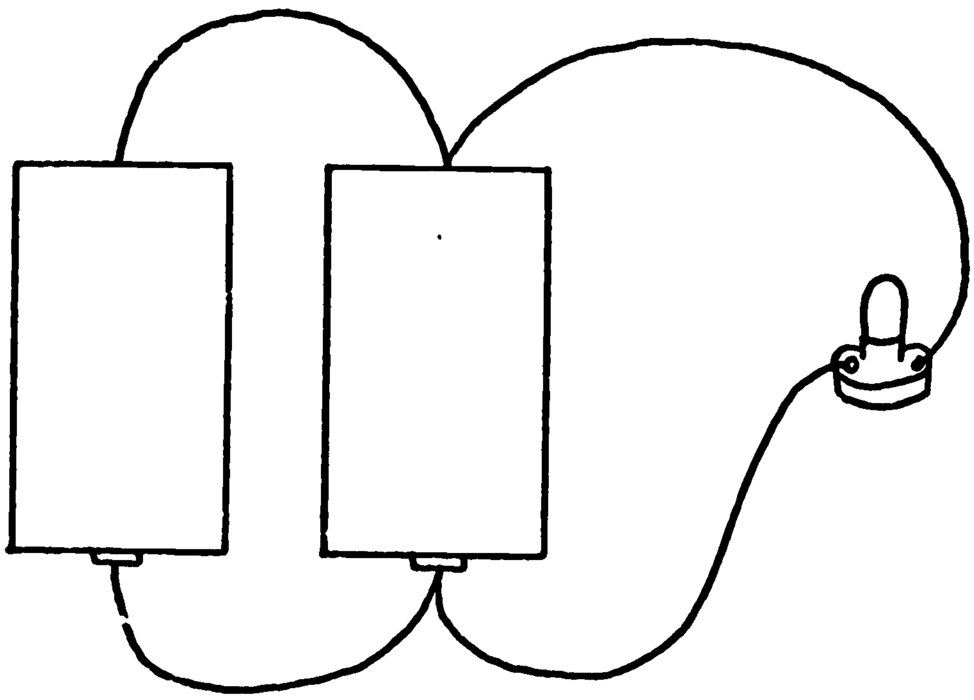
2. THE 3 BATTERIES ADD THEIR POWER TOGETHER.

3. EACH BATTERY GIVES OUT ONLY ONE THIRD ITS ELECTRIC POWER.

4. EACH BATTERY IS ON ITS OWN SEPARATE PATH TO THE BULB.

PINK BULB

THICK COPPER WIRES



IN THIS SET UP:

1. THE 2 BATTERIES ADD THEIR POWER TOGETHER.

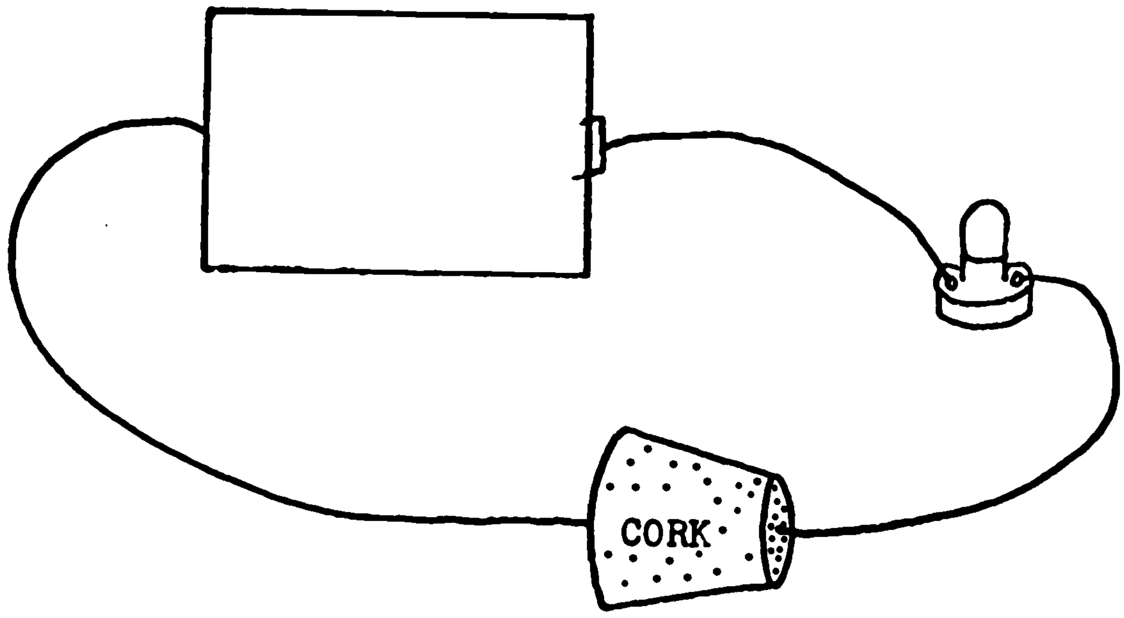
2. EACH BATTERY GIVES OUT ONLY ONE HALF OF ITS ELECTRIC POWER.

3. THE 2 BATTERIES MAKE A SHORT CIRCUIT.

4. THE 2 BATTERIES DON'T WORK BECAUSE THEY ARE POINTING THE WRONG WAY.

PINK BULB

THICK COPPER WIRE



IN THIS SET UP:

1. THE ELECTRIC POWER CAN'T GET THROUGH THE CORK.

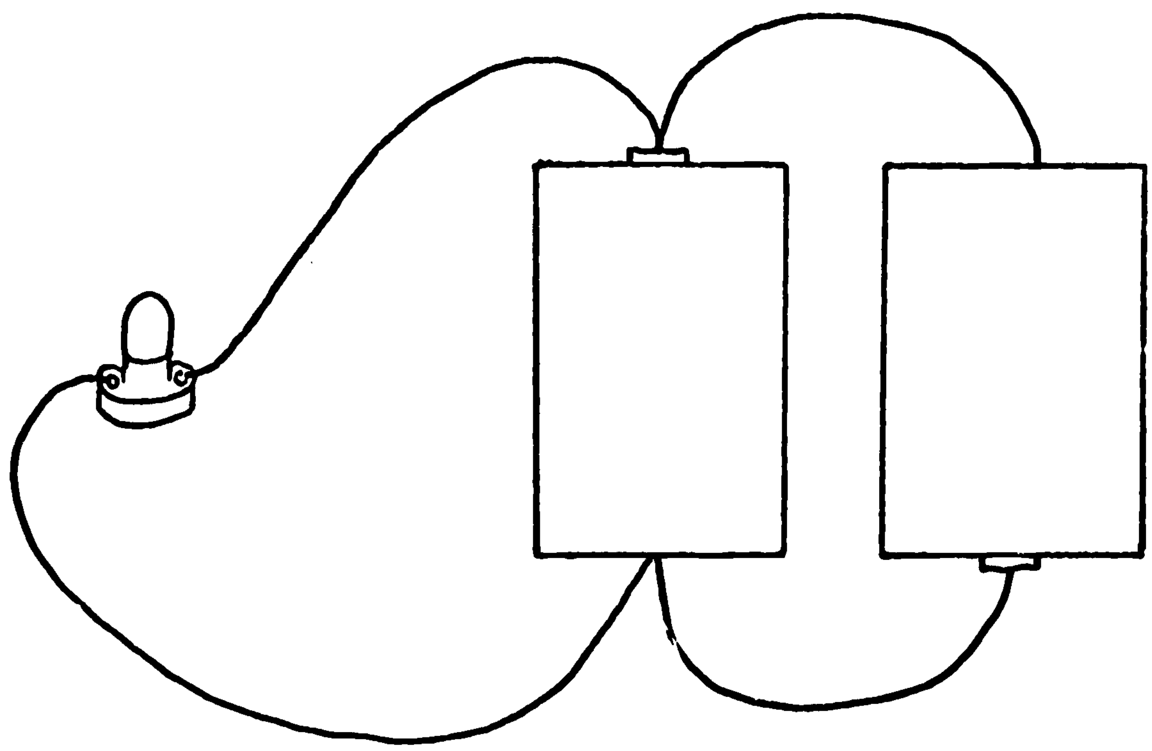
2. THE ELECTRIC POWER CAN TRAVEL THROUGH THE CORK EASILY.

3. THE ELECTRIC POWER CAN'T TRAVEL AT ALL WHEN THE BATTERY IS ON ITS SIDE.

4. THE ELECTRIC POWER HAS A HARD TIME PASSING THROUGH COPPER WIRE.

PINK BULB

THICK COPPER WIRE



IN THIS SET UP:

1. THE 2 BATTERIES BLOW OUT THE BULB.

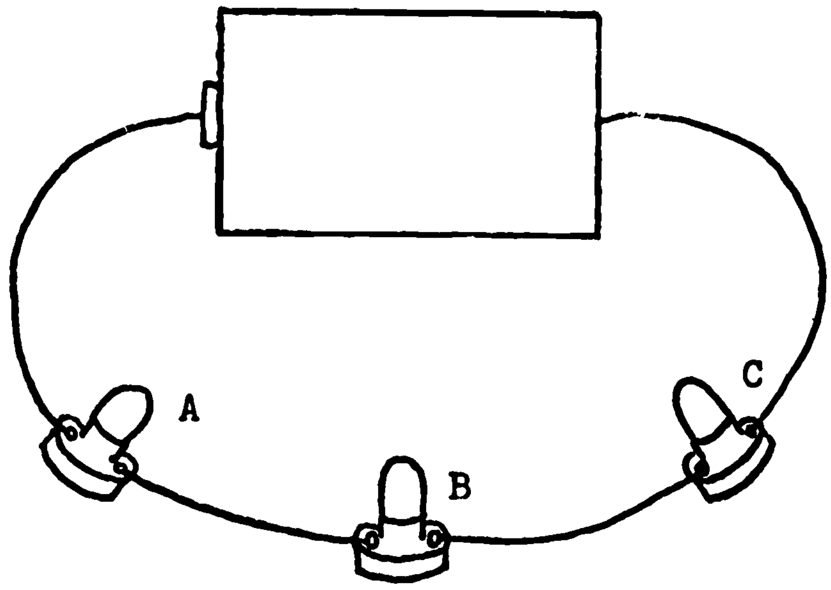
2. EACH BATTERY GIVES OUT ONLY ONE HALF OF ITS ELECTRIC POWER.

3. THE 2 BATTERIES MAKE A SHORT CIRCUIT.

4. THE 2 BATTERIES ADD THEIR POWER TOGETHER.

ALL PINK BULBS

THICK COPPER WIRES



IN THIS SET UP:

1. EACH BULB GETS POWER FROM THE BATTERY AT A SLOWER RATE THAN ONE STANDARD BULB DOES.

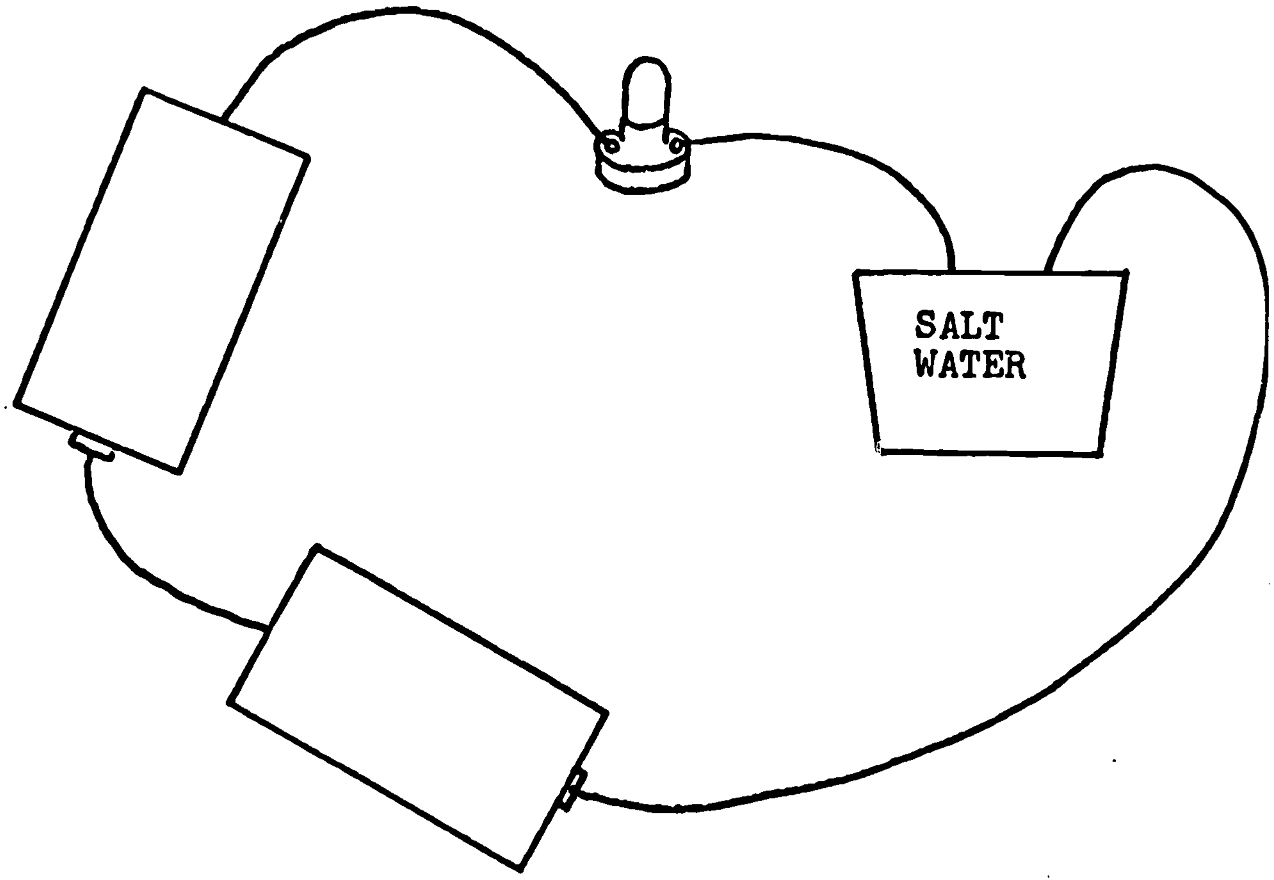
2. BULB A GETS POWER FROM THE BATTERY AT A FASTER RATE THAN BULBS B OR C DO.

3. EACH BULB GETS POWER FROM THE BATTERY AT THE SAME RATE THAT ONE STANDARD BULB DOES.

4. NONE OF THE BULBS GET ANY ELECTRIC POWER BECAUSE THE BATTERY CAN'T WORK ON ITS SIDE.

PINK BULB

THICK COPPER WIRES



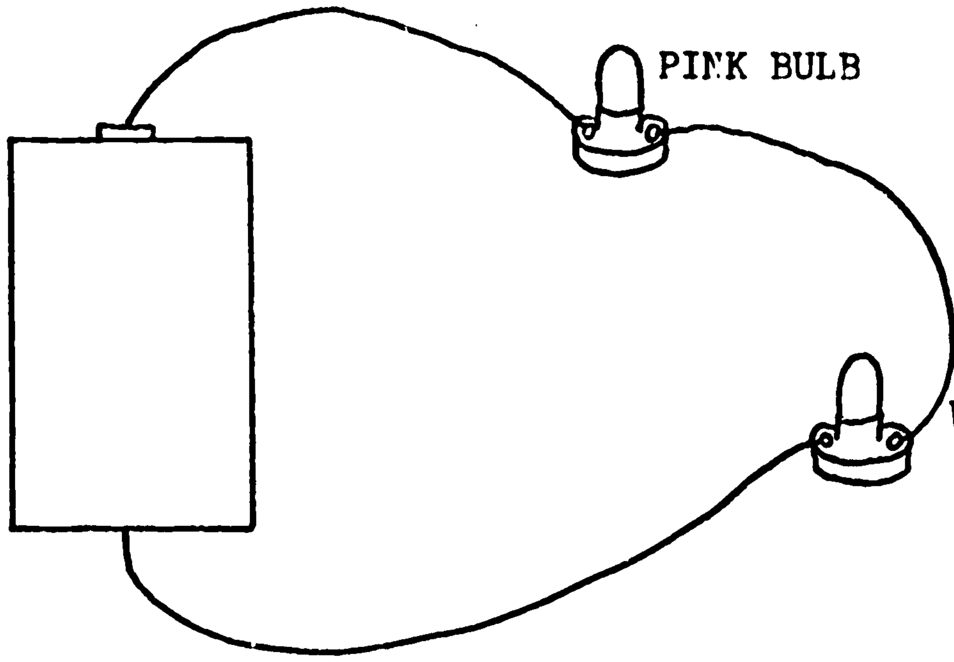
IN THIS SET UP:

1. THE ELECTRIC POWER HAS TROUBLE GETTING THROUGH THE COPPER WIRES.

2. THE ELECTRIC POWER HAS TROUBLE GETTING THROUGH THE SALT WATER.

3. THE ELECTRIC POWER PASSES EASILY THROUGH THE SALT WATER.

4. THE ELECTRIC POWER PASSES EASILY THROUGH THE GLASS OF THE BULB.



IN THIS SET UP:

1. THE PINK BULB AND THE WHITE BULB EACH GET ELECTRIC POWER AT THE SAME RATE THAT ONE STANDARD BULB DOES.

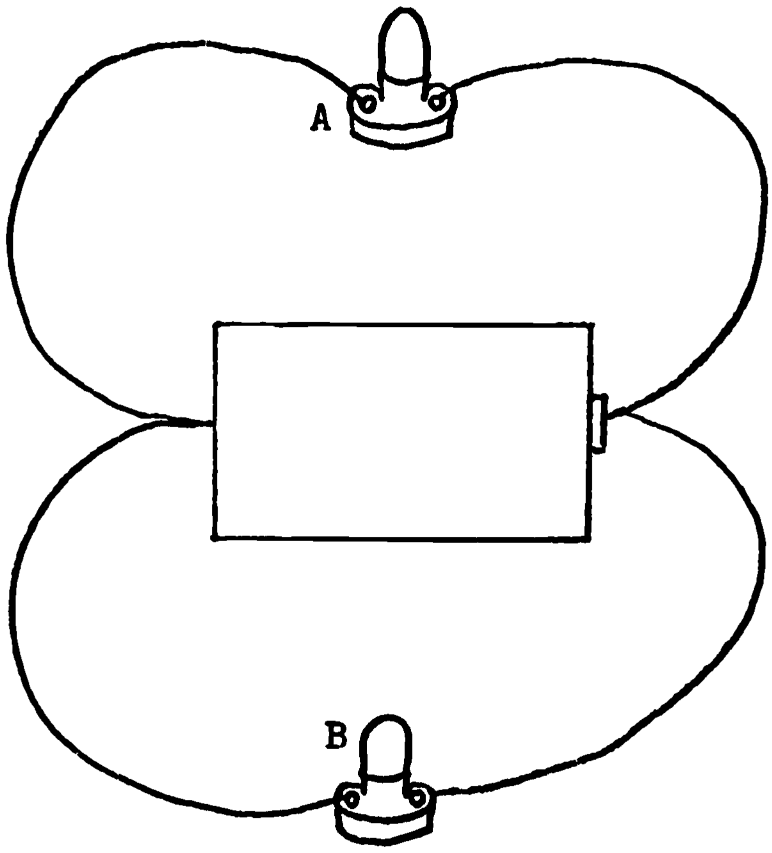
2. THE WHITE BULB SHORT CIRCUITS THE PINK BULB.

3. THE WHITE BULB GETS POWER AT A FASTER RATE THAN THE PINK BULB.

4. THE PINK BULB GETS POWER AT A FASTER RATE THAN THE WHITE BULB.

ALL PINK BULBS

THICK COPPER WIRES



IN THIS SET UP:

1. NEITHER BULB GETS ANY POWER BECAUSE THE BATTERY CAN'T WORK ON ITS SIDE.

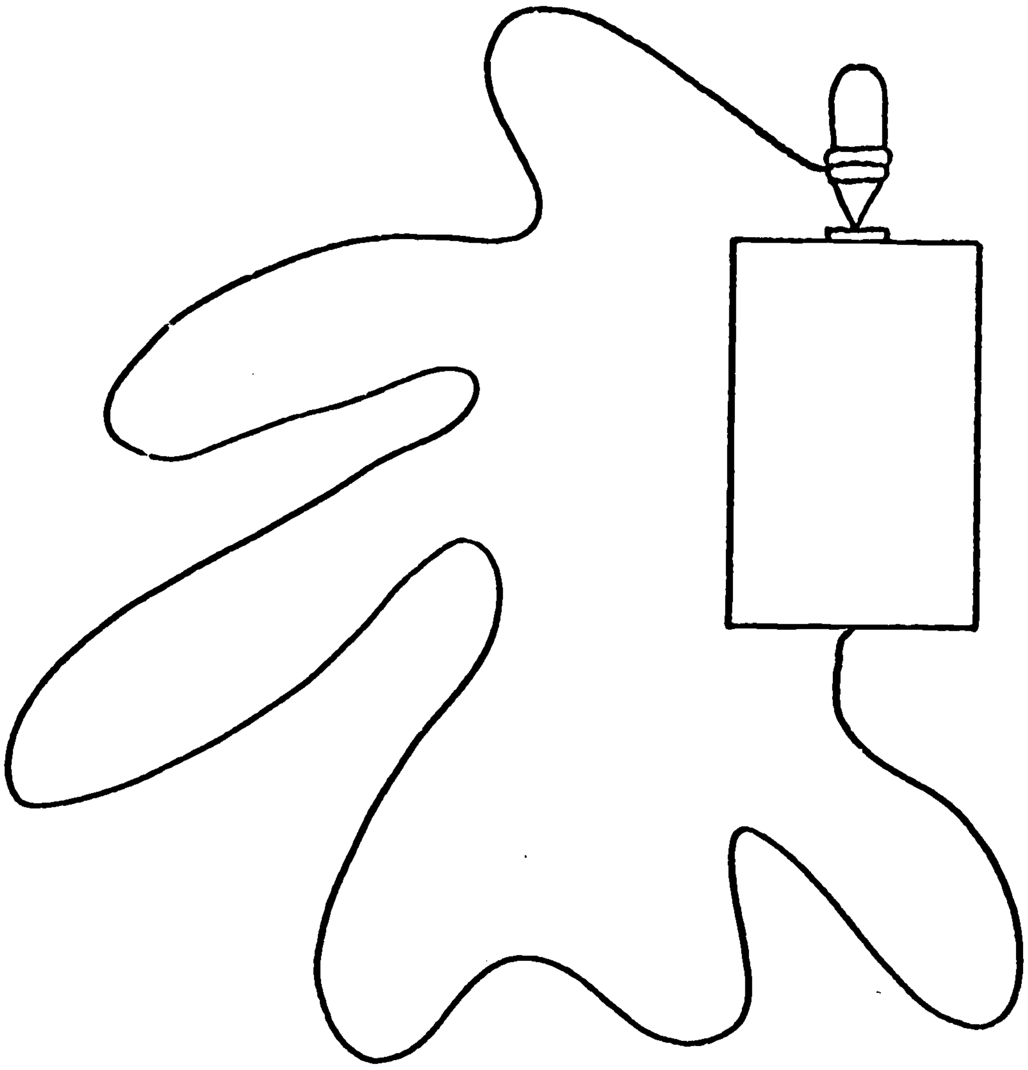
2. NEITHER BULB GETS ANY POWER BECAUSE ONE BULB SHORT CIRCUITS THE OTHER.

3. EACH BULB GETS ELECTRIC POWER AT THE SAME RATE AS ONE STANDARD BULB DOES.

4. EACH BULB GETS ELECTRIC POWER AT A SLOWER RATE THAN ONE STANDARD BULB DOES.

PINK BULB

THICK COPPER WIRE



IN THIS SET UP:

1. THE ELECTRIC POWER GOES THROUGH THE WIRE EASILY BECAUSE IT IS COPPER.

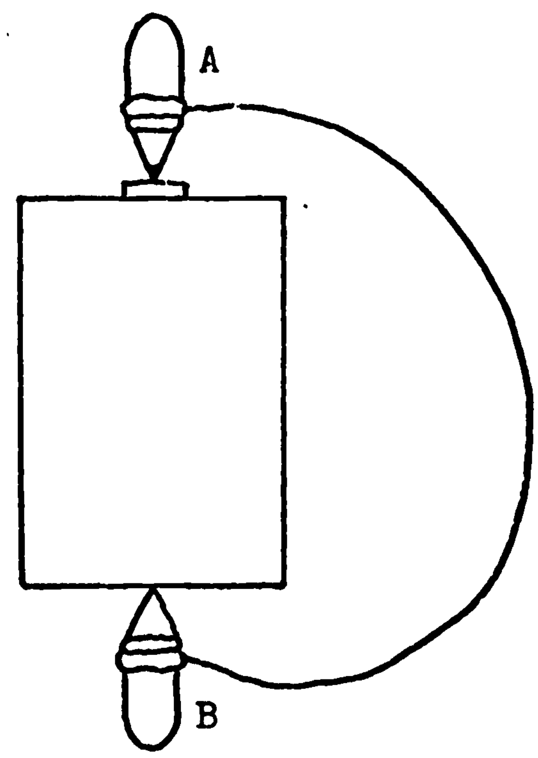
2. THE ELECTRIC POWER GOES THROUGH THE WIRE EASILY BECAUSE IT IS LONG.

3. THE ELECTRIC POWER HAS TROUBLE GETTING THROUGH THE WIRE BECAUSE IT IS COPPER.

4. THE ELECTRIC POWER HAS TROUBLE GETTING THROUGH THE WIRE BECAUSE IT IS LONG.

ALL PINK BULBS

THICK COPPER WIRE



IN THIS SET UP:

1. BULB B CAN'T WORK BECAUSE IT IS NOT TOUCHING THE SILVER CAP OF THE BATTERY.

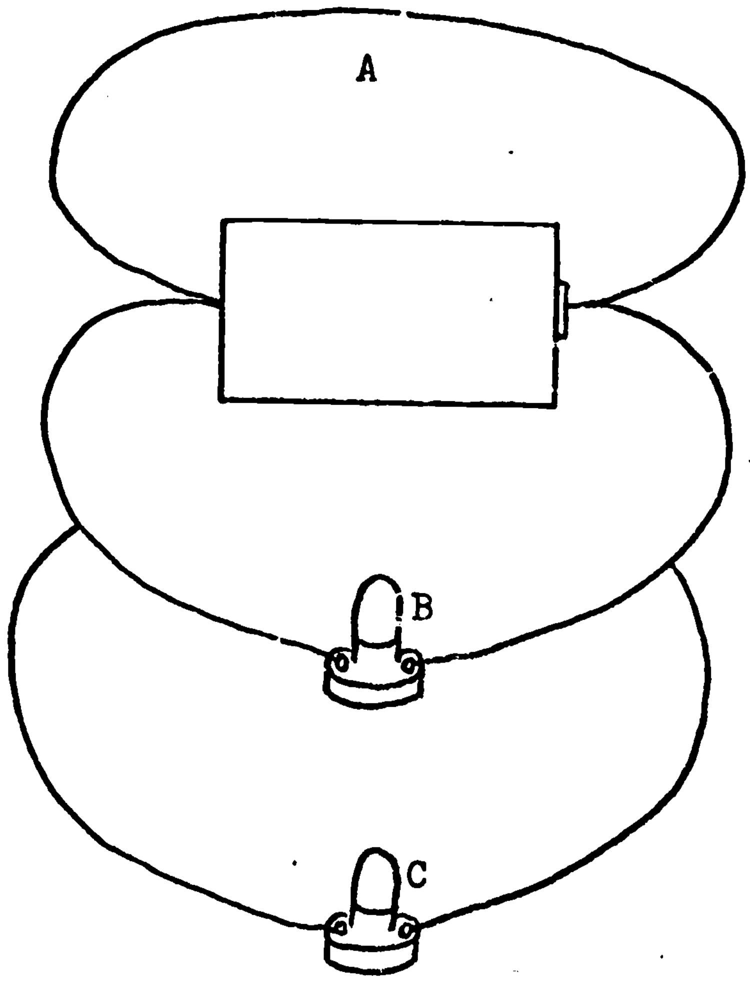
2. NEITHER BULB CAN WORK BECAUSE EACH SHORT CIRCUITS THE OTHER.

3. NEITHER BULB CAN WORK BECAUSE THE BATTERY CAN'T SEND OUT POWER IN TWO DIRECTIONS AT ONCE.

4. BOTH BULBS WORK BECAUSE EACH COMPLETES THE CIRCUIT FOR THE OTHER.

ALL PINK BULBS

THICK COPPER WIRES



IN THIS SET UP:

1. WIRE A PUMPS EXTRA POWER TO THE BULBS.

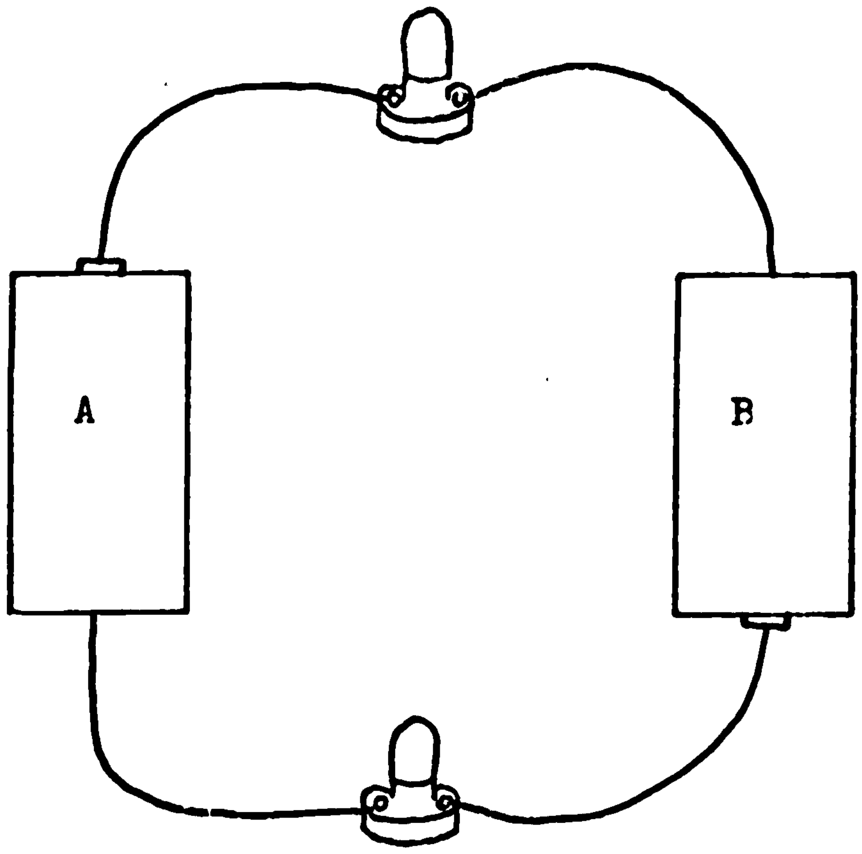
2. WIRE A SHORT CIRCUITS THE BULBS.

3. BULB B SHORT CIRCUITS BULB C.

4. THERE ARE TOO MANY WIRES FOR THE ELECTRIC POWER TO TRAVEL AT ALL.

PINK BULBS

THICK COPPER WIRES



IN THIS SET UP:

1. THE 2 BATTERIES ADD THEIR POWER TOGETHER.

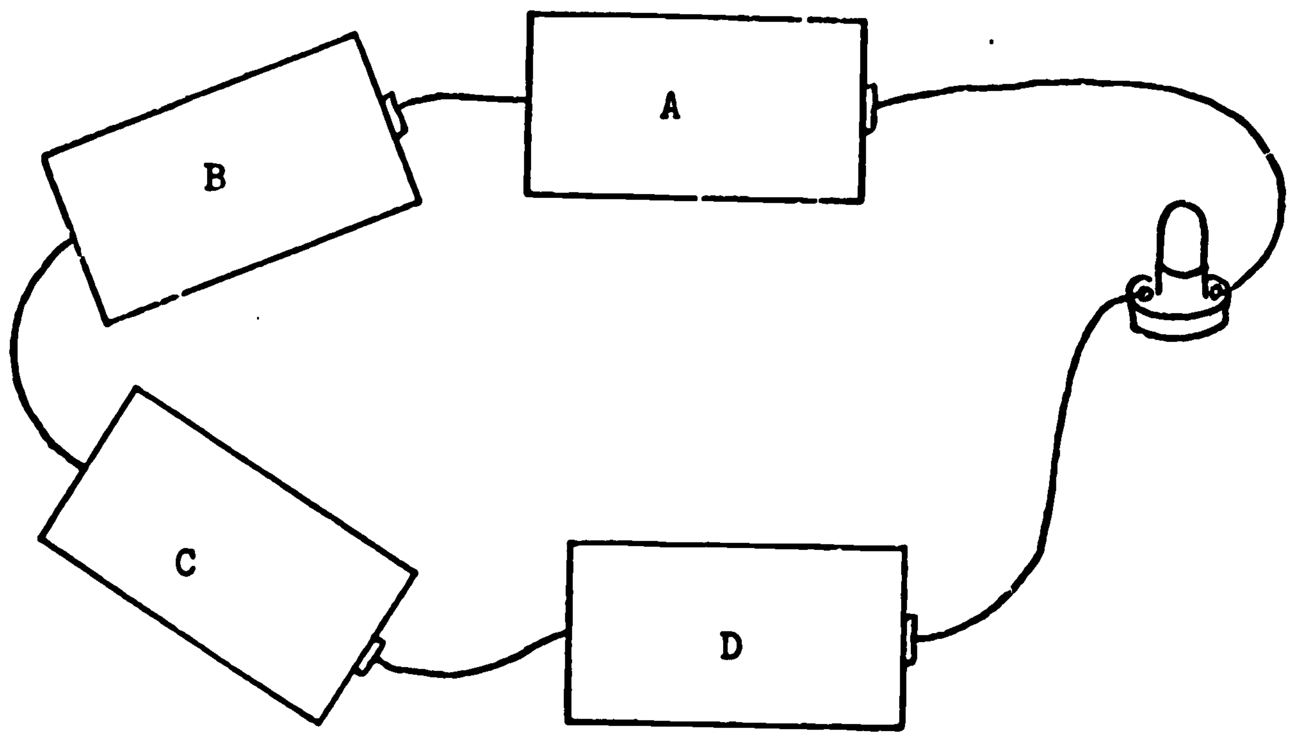
2. BATTERY A CANCELS OUT BATTERY B.

3. BATTERY B BREAKS THE CIRCUIT BECAUSE IT IS UPSIDE DOWN.

4. EACH BATTERY GIVES OUT ONLY ONE HALF OF ITS ELECTRIC POWER.

PINK BULB

THICK COPPER WIRES



IN THIS SET UP:

1. EACH BATTERY GIVES OUT ONLY ONE FOURTH OF ITS ELECTRIC POWER.

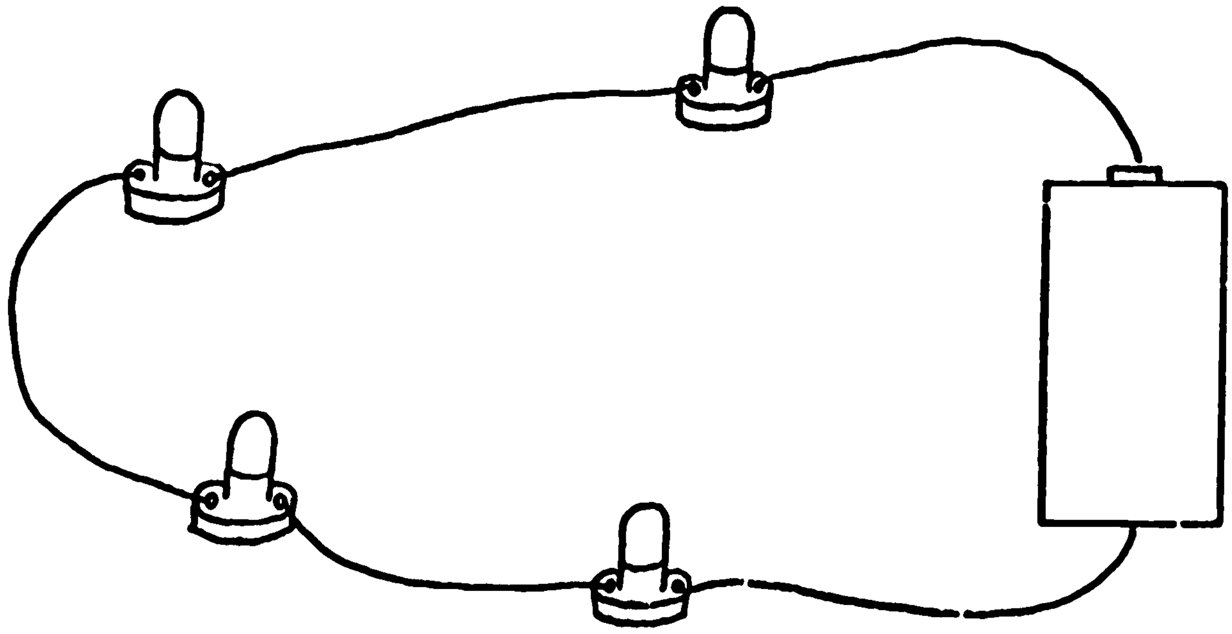
2. BATTERIES B AND C CANCEL OUT EACH OTHER AND BATTERIES A AND D CANCEL OUT EACH OTHER.

3. ALL 4 BATTERIES ADD THEIR POWER TOGETHER.

4. BATTERIES A AND B ARE NOT USED UNTIL BATTERIES C AND D RUN DOWN.

ALL PINK BULBS

THICK COPPER WIRES



IN THIS SET UP:

1. EACH BULB GETS ELECTRIC POWER AT A MUCH FASTER RATE THAN ONE STANDARD BULB DOES.

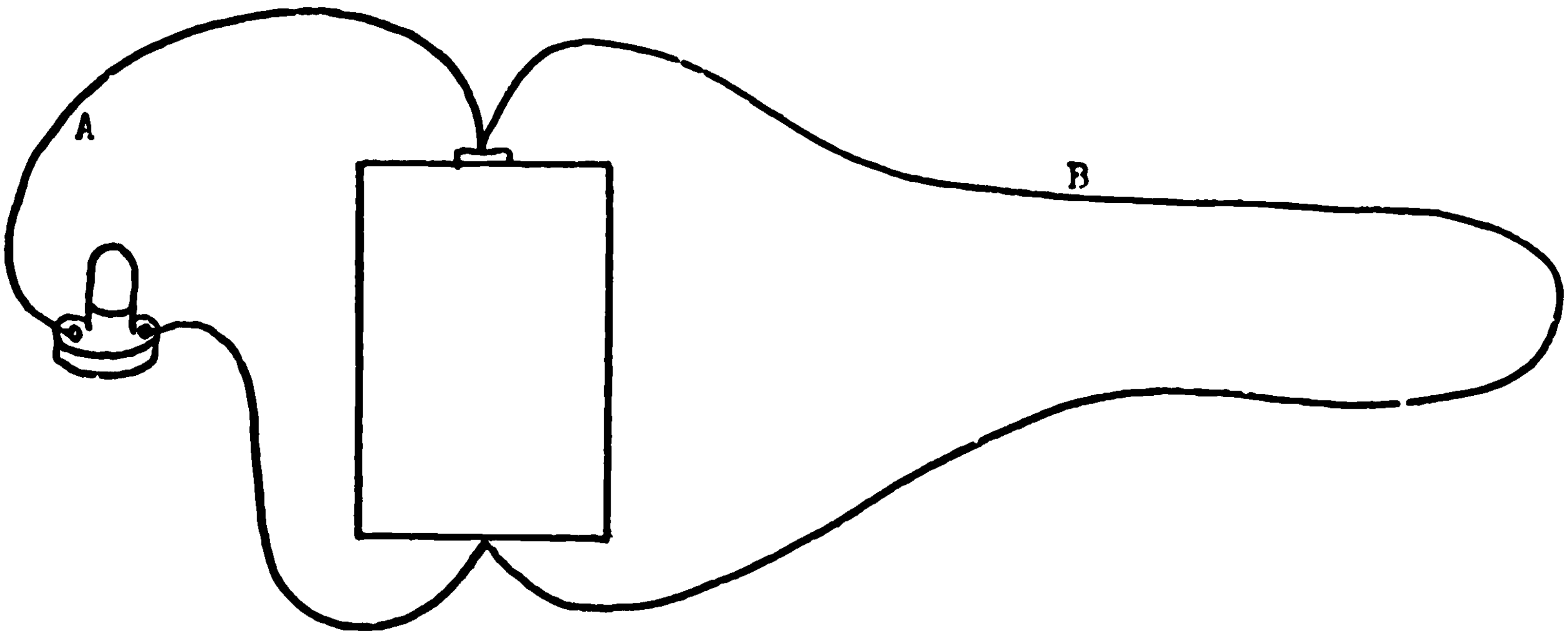
2. ALL 4 BULBS BLOW OUT THE BATTERY.

3. EACH BULB GETS ELECTRIC POWER AT THE SAME RATE THAT ONE STANDARD BULB DOES.

4. EACH BULB GETS ELECTRIC POWER AT A MUCH SLOWER RATE THAN ONE STANDARD BULB DOES.

PINK BULB

THICK COPPER WIRES



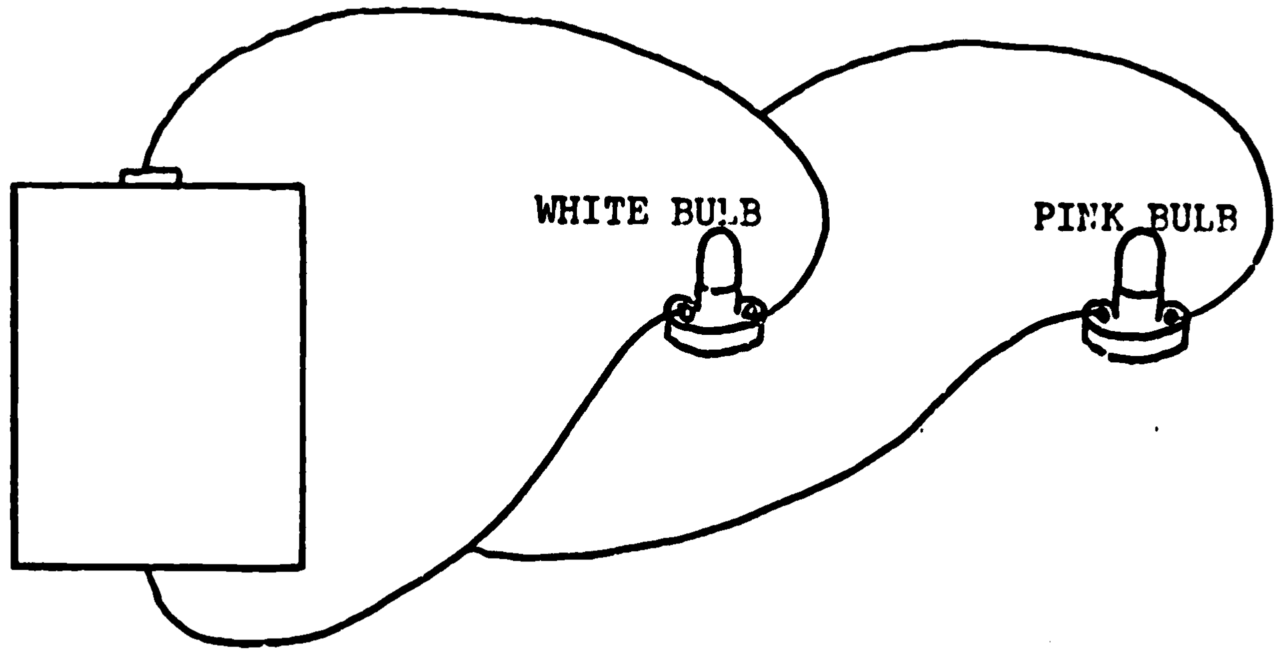
IN THIS SET UP:

1. THE ELECTRIC POWER TAKES THE EASIER PATH B BECAUSE IT IS PLAIN COPPER.

2. THE ELECTRIC POWER TAKES THE EASIER PATH A BECAUSE PATH B IS TOO LONG.

3. THE ELECTRIC POWER TAKES THE EASIER PATH B BECAUSE PATH A HAS TOO MANY WIRES.

4. THE ELECTRIC POWER TAKES THE EASIER PATH A BECAUSE PATH B IS NOT THIN ENOUGH.



IN THIS SET UP:

1. THE PINK BULB AND THE WHITE BULB EACH GET ELECTRIC POWER AT THE SAME RATE THAT ONE STANDARD BULB DOES.

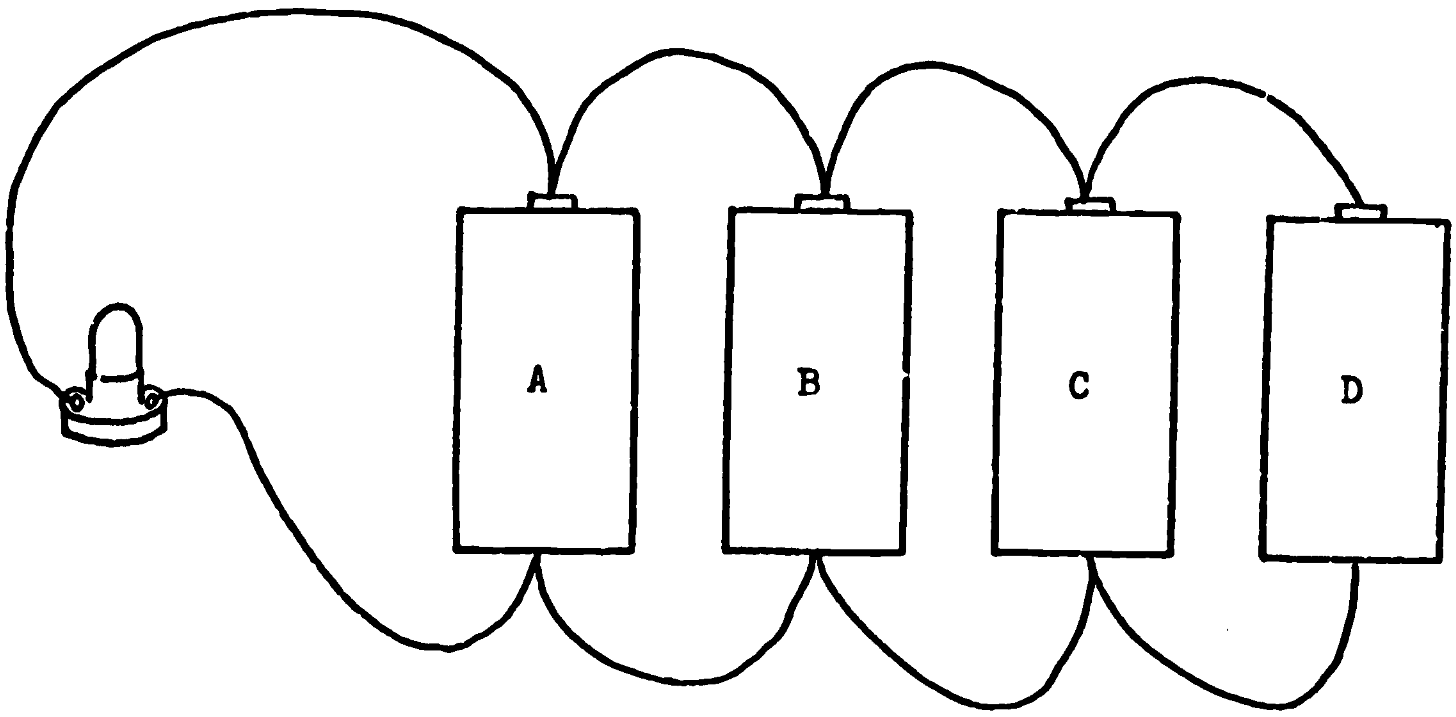
2. THE PINK BULB AND THE WHITE BULB EACH GETS ELECTRIC POWER AT A SLOWER RATE THAN ONE STANDARD BULB DOES.

3. THE WHITE BULB GETS POWER AT A FASTER RATE THAN THE PINK BULB.

4. THE PINK BULB GETS POWER AT A FASTER RATE THAN THE WHITE BULB.

PINK BULB

THICK COPPER WIRES



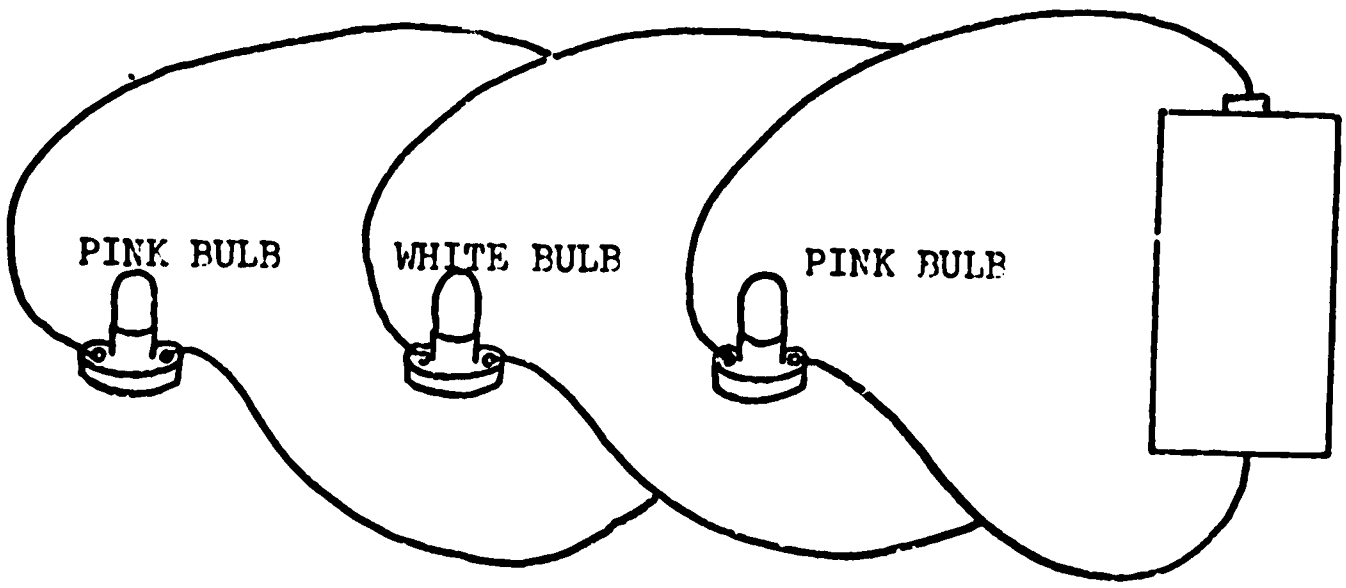
IN THIS SET UP:

1. BATTERIES A AND B CANCEL EACH OTHER OUT, AND BATTERIES C AND D CANCEL EACH OTHER OUT.

2. BATTERIES B AND C MAKE A SHORT CIRCUIT.

3. EACH BATTERY GIVES OUT ONLY ONE FOURTH OF ITS ELECTRIC POWER.

4. ALL 4 BATTERIES ADD THEIR POWER TOGETHER.



IN THIS SET UP:

1. THE PINK BULBS EACH GET ELECTRIC POWER AT A FASTER RATE THAN THE WHITE BULB DOES.

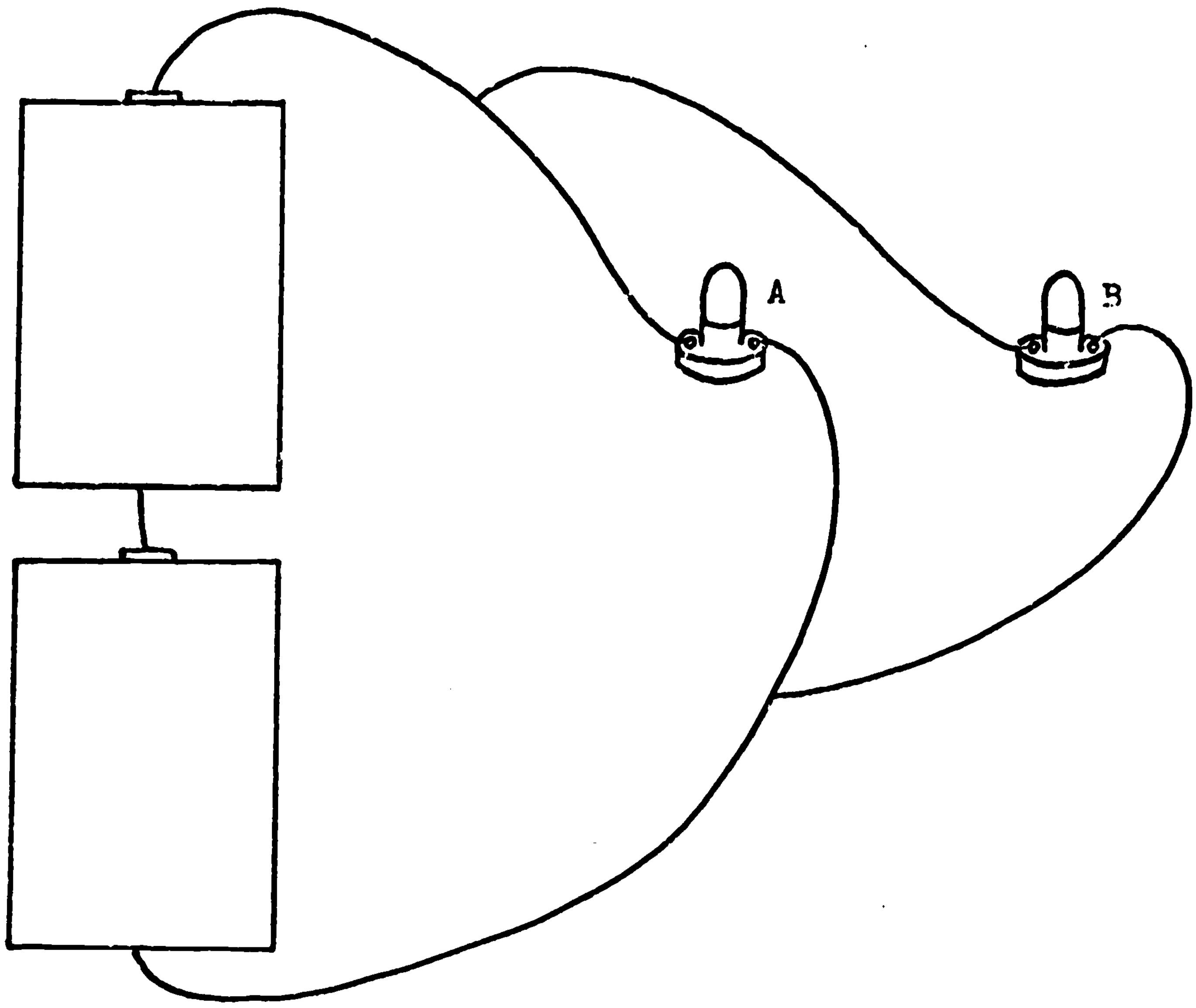
2. THE WHITE BULB GETS ELECTRIC POWER AT A FASTER RATE THAN THE PINK BULBS DO.

3. EACH BULB GETS THE ELECTRIC POWER AT A SLOWER RATE THAN ONE STANDARD BULB DOES.

4. EACH BULB GETS THE ELECTRIC POWER AT THE SAME RATE THAT ONE STANDARD BULB DOES.

ALL PINK BULBS

THICK COPPER WIRES



IN THIS SET UP:

1. BULBS A AND B EACH GET ELECTRIC POWER AT A FASTER RATE THAN ONE STANDARD BULB DOES.

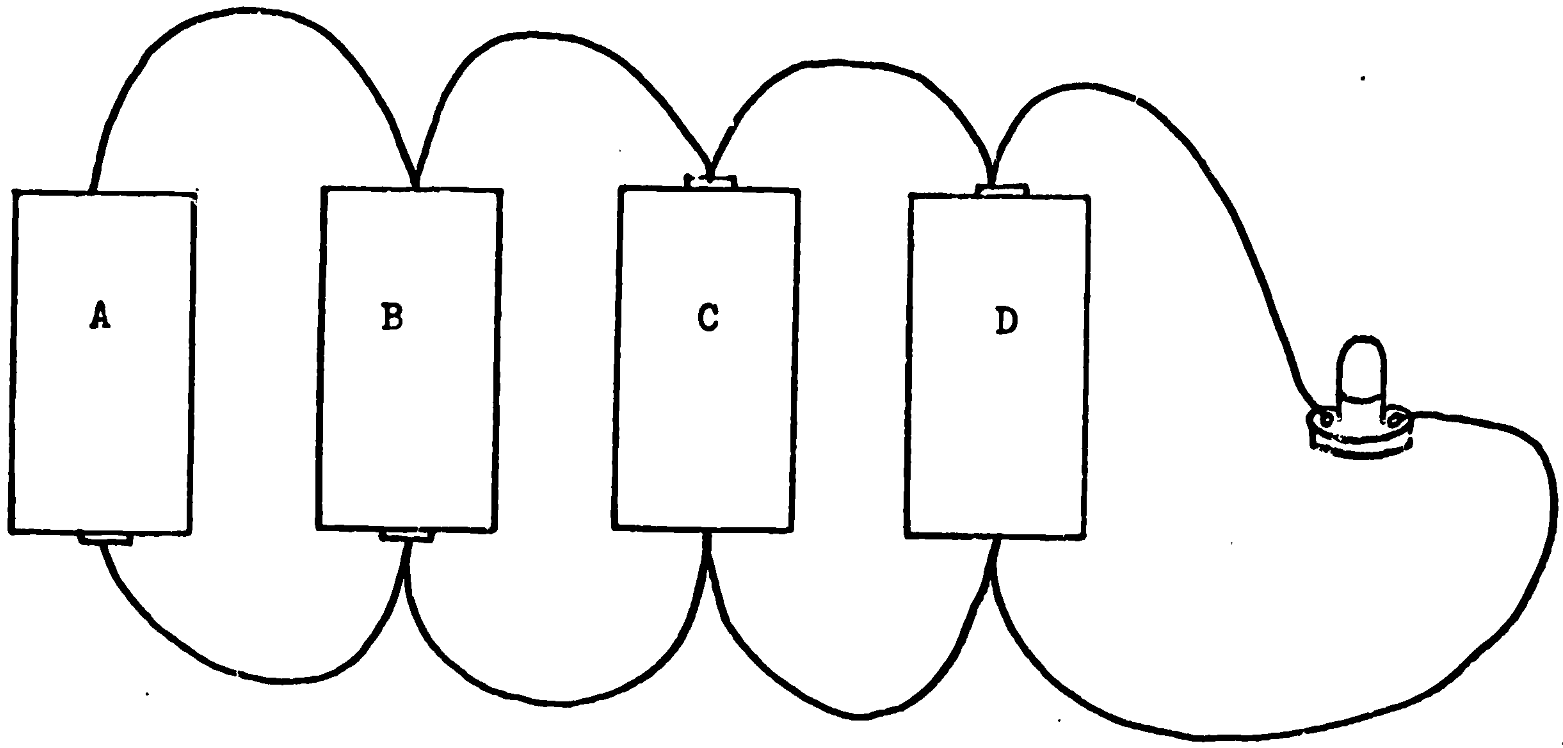
2. BULBS A AND B EACH GET ELECTRIC POWER AT THE SAME RATE THAT ONE STANDARD BULB DOES.

3. BULBS A AND B GET NO ELECTRIC POWER BECAUSE THE BATTERIES CANCEL EACH OTHER OUT.

4. BULBS A AND B GET NO ELECTRIC POWER BECAUSE THERE ARE TOO MANY WIRES.

PINK BULB

THICK COPPER WIRES



IN THIS SET UP:

1. BATTERIES B AND C MAKE A SHORT CIRCUIT.

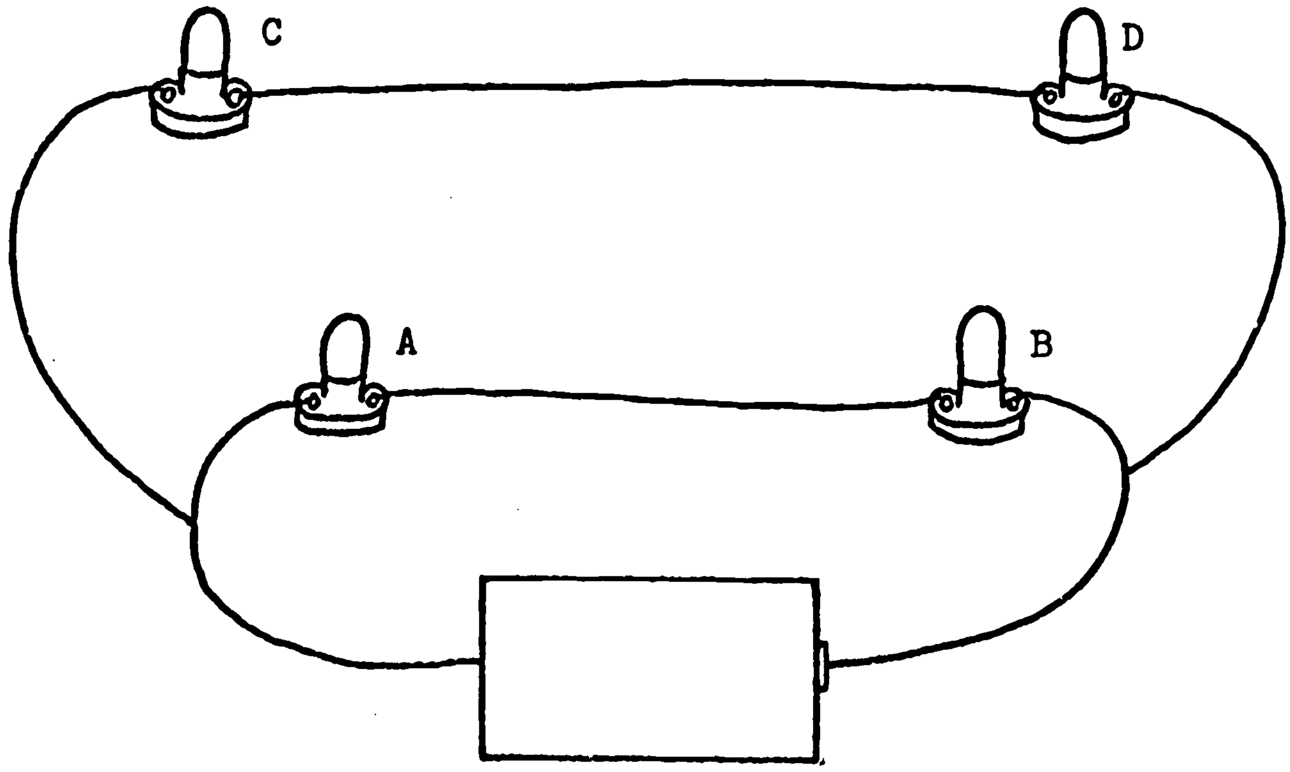
2. BATTERIES A AND B CANCEL EACH OTHER OUT AND BATTERIES C AND D CANCEL EACH OTHER OUT.

3. ALL 4 BATTERIES ADD THEIR ELECTRIC POWER TOGETHER.

4. EACH BATTERY GIVES OUT ONLY ONE FOURTH OF ITS ELECTRIC POWER.

ALL PINK BULBS

THICK COPPER WIRES



IN THIS SET UP:

1. EACH BULB GETS THE ELECTRIC POWER AT THE SAME RATE THAT ONE STANDARD BULB DOES.

2. EACH BULB GETS THE ELECTRIC POWER AT A SLOWER RATE THAN ONE STANDARD BULB DOES.

3. EACH BULB GETS THE ELECTRIC POWER FOUR TIMES AS QUICKLY AS A STANDARD BULB DOES.

4. BULBS A AND B EACH GET ELECTRIC POWER AT A FASTER RATE THAN BULBS C AND D DO BECAUSE THEY ARE CLOSER TO THE BATTERY.